GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:18:09 ; Search time 65.5868 Seconds
(without alignments)
4734.482 Million cell updates/sec

Title: US-10-658-782-6
Perfect score: 5912
Sequence: 1 MATKAVCVLKGDGPVQGIIN.....GNKDRRSTGKSWGKPGYPWP 1099

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5912	100.0	1099	5	AAU76378	Aau76378 HCV multi
2	5912	100.0	1099	6	ABG72262	Abg72262 HCV multi
3	4032	68.2	829	5	AAE18690	Aae18690 Multiple
4	4032	68.2	829	7	ADC06769	Adc06769 Chimeric
5	3829.5	64.8	1021	2	AAW34481	Aaw34481 HCV antig
6	3829.5	64.8	1021	2	AAW40039	Aaw40039 Fusion pr
7	3829.5	64.8	1021	5	AAE22050	Aae22050 pSOD/c200
8	3050.5	51.6	841	2	AAR68547	Aar68547 HCV prote
9	3050.5	51.6	841	6	ABO27020	Abo27020 Hepatitis
10	3050.5	51.6	841	7	ADA07875	Ada07875 HCV prote
11	3047.5	51.5	841	2	AAW01701	Aaw01701 hSOD-HCV
12	3047.5	51.5	841	2	AAW46397	Aaw46397 Amino aci
13	3047.5	51.5	841	2	AAW97609	Aaw97609 Amino aci
14	3042.5	51.5	840	2	AAR14349	Aar14349 HCV prote
15	2909.5	49.2	2261	1	AAP90164	Aap90164 Peptide e
16	2909.5	49.2	2436	1	AAP92050	Aap92050 Sequence
17	2909.5	49.2	2436	1	AAP90288	Aap90288 Peptide e
18	2909.5	49.2	2772	3	AAB18540	Aab18540 Protein e
19	2909.5	49.2	2955	2	AAY14975	Aay14975 Amino aci
20	2909.5	49.2	2955	3	AAB18541	Aab18541 Polyprote
21	2909.5	49.2	3011	2	AAR90931	Aar90931 Hepatitis
22	2909.5	49.2	3011	2	AAW34480	Aaw34480 HCV polyp
23	2909.5	49.2	3011	2	AAW40038	Aaw40038 HCV polyp
24	2909.5	49.2	3011	5	AAE22049	Aae22049 Hepatitis
25	2906.5	49.2	2301	1	AAP92047	Aap92047 Sequence
26	2906.5	49.2	2772	2	AAR08123	Aar08123 Hepatitis
27	2900.5	49.1	3011	2	AAR21519	Aar21519 Compiled
28	2900.5	49.1	3011	5	AAU84597	Aau84597 HCV polyp
29	2897.5	49.0	2435	2	AAR25135	Aar25135 HCV polyp
30	2896.5	49.0	1766	1	AAP92041	Aap92041 Hepatitis
31	2895.5	49.0	2894	2	AAR24440	Aar24440 Composite
32	2893.5	48.9	2816	2	AAR34009	Aar34009 HCV-1 pol
33	2891.5	48.9	3011	2	AAR31621	Aar31621 Hepatitis
34	2890.5	48.9	1786	1	AAP90158	Aap90158 Protein s
35	2890.5	48.9	2894	2	AAR70230	Aar70230 Composite
36	2883.5	48.8	2436	2	AAR28582	Aar28582 HCV amino
37	2880.5	48.7	2955	2	AAR08124	Aar08124 Hepatitis
38	2877.5	48.7	3011	5	AAE22052	Aae22052 Hepatitis
39	2875	48.6	3011	2	AAW77397	Aaw77397 Hepatitis
40	2875	48.6	3011	6	ABP71460	Abp71460 Amino aci
41	2875	48.6	3012	5	AAU99289	Aau99289 Hepatitis
42	2875	48.6	3012	6	ABU61848	Abu61848 HCV H77 c
43	2868	48.5	2984	4	AAE00442	Aae00442 Hepatitis
44	2868	48.5	3011	2	AAW77398	Aaw77398 Hepatitis
45	2868	48.5	3011	2	AAW98021	Aaw98021 Infectiou

ALIGNMENTS

RESULT 1
AAU76378
ID AAU76378 standard; protein; 1099 AA.
XX
AC AAU76378;
XX
DT 08-MAY-2002 (first entry)
XX
DE HCV multiple epitope fusion antigen (MEFA) 7.1 protein sequence.
XX
KW Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion;
KW immunoassay solid support; multiple epitope fusion antigen; MEFA;
KW non-structural protein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
PN WO200196870-A2.
XX
PD 20-DEC-2001.
XX
PF 14-JUN-2001; 2001WO-US019156.
XX
PR 15-JUN-2000; 2000US-0212082P.
PR 02-APR-2001; 2001US-0280811P.
PR 02-APR-2001; 2001US-0280867P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Chien DY, Arcangel P, Tandeske L, George-Nasciemento C, Coit D;
PI Medina-Selby A;
XX
DR WPI; 2002-090228/12.
DR N-PSDB; ABK15345.
XX
PT Immunoassay solid support, useful for detecting hepatitis C virus
PT infection in biological sample, comprises HCV NS3/4a conformational
PT epitope and multiple epitope fusion antigen bound to the support.
XX
PS Claim 5; Fig 5; 92pp; English.
XX
CC The present invention relates to a new immunoassay solid support
CC consisting essentially of at least one hepatitis C virus (HCV) NS3/4a
CC conformational epitope and a multiple epitope fusion antigen (MEFA),
CC bound to the support. The NS3/4a conformational epitope and/or MEFA
CC reacts specifically with anti-HCV antibodies present in a biological
CC sample from an HCV-infected individual. The immunoassay of the invention
CC is useful for detecting hepatitis C virus infection in a biological

CC sample. The method of the invention provides a sensitive, accurate
CC diagnostic and prognostic tool to provide adequate patient care and to
CC prevent transmission of HCV by blood and by blood products, or by
CC personal contact. Use of NS3/4a conformational epitope in combination
CC with MEFA, provides a sensitive and reliable method for detecting early
CC HCV seroconversion. Use of MEFA has the added advantages of decreasing
CC masking problems, improving sensitivity in detecting antibodies by
CC allowing a greater number of epitopes on a unit surface area of
CC substrate, and improving substrate. Detection accuracy is increased and
CC the incidence of false results is reduced because of the identification
CC and the use of highly immunogenic HCV antigens which are present during
CC the early stages of HCV seroconversion. The present amino acid sequence
CC represents the multiple epitope fusion antigen (MEFA) 7.1 of the
CC invention
XX
SQ Sequence 1099 AA;

Query Match 100.0%; Score 5912; DB 5; Length 1099;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1099; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180

Qy        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240

Qy        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300

Qy        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360

Qy        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420

Qy        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480

Qy        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540

Qy        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600

Qy        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660

Qy        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSACSGKPAIIPDREVLYREFDEMEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSACSGKPAIIPDREVLYREFDEMEE 720

Qy        721 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 780

Qy        781 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 840

Qy        841 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 900

Qy        901 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 960

Qy        961 KRNTNRRPQDVKFPGGGQIVGRRGPPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSW 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 KRNTNRRPQDVKFPGGGQIVGRRGPPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSW 1020

Qy       1021 GKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGRRGPPIPKARRPEGRTWAQPGYPWPLYG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGRRGPPIPKARRPEGRTWAQPGYPWPLYG 1080

Qy       1081 NKDRRSTGKSWGKPGYPWP 1099
              |||||||||||||||||||
Db       1081 NKDRRSTGKSWGKPGYPWP 1099
```

RESULT 2
ABG72262
ID ABG72262 standard; protein; 1099 AA.
XX
AC ABG72262;
XX
DT 06-MAR-2003 (first entry)
XX
DE HCV multiple epitope fusion antigen 7.1 (MEFA 7.1).
XX
KW Immunoassay solid support; Hepatitis C Virus type-1; HCV-1; HCV-2;
KW NS3/4a conformational epitope; multiple epitope fusion antigen 7.1;
KW MEFA 7.1; anti-HCV antibody; NS3/4a conformational antigen; HCV-3;
KW HCV infection; Hepatitis C Virus type-2; Hepatitis C Virus type-3;
KW mutant; mutein.
XX
OS Hepatitis C virus type 1.
OS Hepatitis C virus type 2.
OS Hepatitis C virus type 3.
OS Synthetic.
OS Chimeric.
XX

FH	Key	Location/Qualifiers
FT	Region	1. .156
FT		/note= "Correspond to amino acids 1-156 of HCV-1 hSOD
FT		superoxide dismutase)"
FT	Region	159. .176
FT		/note= "Correspond to amino acids 303-320 of HCV-1 E1"
FT	Region	179. .199
FT		/note= "Correspond to consensus sequence of amino acids
FT		390-410 of HCV-1 E2 HVR"
FT	Region	200. .230
FT		/note= "Correspond to consensus sequence of amino acids
FT		384-414 of HCV-1 and HCV-2 E2 HVR"
FT	Region	231. .696
FT		/note= "Correspond to amino acids 1193-1658 of HCV-1
FT		helicase"
FT	Region	699. .745
FT		/note= "Correspond to amino acids 1689-1735 of HCV-1 5-1-
FT		1 epitope"
FT	Region	748. .794
FT		/note= "Correspond to amino acids 1689-1735 of HCV-3 5-1-
FT		1 epitope"
FT	Region	797. .843
FT		/note= "Correspond to amino acids 1689-1735 of HCV-2 5-1-
FT		1 epitope"
FT	Region	846. .881
FT		/note= "Correspond to amino acids 1901-1936 of HCV-1

```
FT                   polypeptide C100"
FT   Region          884. .919
FT                   /note= "Correspond to amino acids 2278-2313 of HCV-1 NS5
FT                   region"
FT   Region          922. .957
FT                   /note= "Correspond to amino acids 2278-2313 of HCV-1 NS5
FT                   region"
FT   Region          958. .1028
FT                   /note= "Correspond to core region antigenic determinants
FT                   from amino acids 9-32, 39-42 and 64-88 of HCV-1 and amino
FT                   acids 67-84 of HCV-2"
FT   Region          1029. .1099
FT                   /note= "Correspond to core region antigenic determinants
FT                   from amino acids 9-32, 39-42 and 64-88 of HCV-1 and amino
FT                   acids 67-84 of HCV-2"
XX
PN   US2002146685-A1.
XX
PD   10-OCT-2002.
XX
PF   14-JUN-2001; 2001US-00881654.
XX
PR   15-JUN-2000; 2000US-0212082P.
PR   02-APR-2001; 2001US-0280811P.
PR   02-APR-2001; 2001US-0280867P.
XX
PA   (CHIE/) CHIEN D Y.
PA   (ARCA/) ARCANGEL P.
PA   (TAND/) TANDESKE L.
PA   (GEOR/) GEORGE-NASCIMENTO C.
PA   (COIT/) COIT D.
PA   (MEDI/) MEDINA-SELBY A.
XX
PI   Chien DY,  Arcangel P,  Tandeske L,  George-Nascimento C,  Coit D;
PI   Medina-Selby A;
XX
DR   WPI; 2003-147573/14.
DR   N-PSDB; ABX14411.
XX
PT   Immunoassay solid support for detecting Hepatitis C Virus infection in
PT   biological samples, comprises Hepatitis C Virus conformational epitope
PT   and multiple epitope fusion antigen.
XX
PS   Claim 25; Fig 5A-5F; 45pp; English.
XX
CC   The present invention relates to immunoassays comprising Hepatitis C
CC   Virus (HCV) NS3/4a conformational epitope and multiple epitope fusion
CC   antigen (MEFA), bound to a solid support. The NS3/4a epitope and/or the
CC   multiple epitope fusion antigen react with anti-HCV antibodies present in
CC   a biological sample from an HCV-infected individual. The immunoassays and
CC   methods of the invention are useful for detecting HCV infection in a
CC   biological sample. The inventive immunoassay solid support provides a
CC   sensitive and reliable method for detecting early HCV seroconversion. The
CC   assays can detect HCV infection caused by any six known genotypes of HCV.
CC   The use of the multiple epitope fusion proteins decreases masking
CC   problems, improves sensitivity in detecting antibodies by allowing a
CC   greater number of epitopes on a unit area of substrate, and improves
CC   selectivity. The present sequence represents HCV multiple epitope fusion
CC   antigen 7.1 (MEFA 7.1), a mutant HCV polyprotein derived from various
CC   regions of HCV type 1, 2, or 3 (HCV-1, HCV-2, or HCV-3) polypeptide
CC   sequences
XX
SQ   Sequence 1099 AA;

  Query Match             100.0%;  Score 5912;  DB 6;  Length 1099;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1099;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180

Qy        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240

Qy        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300

Qy        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360

Qy        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420

Qy        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480

Qy        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540

Qy        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600

Qy        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660

Qy        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSACSGKPAIIPDREVLYREFDEMEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSACSGKPAIIPDREVLYREFDEMEE 720

Qy        721 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 780

Qy        781 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 840

Qy        841 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 900

Qy        901 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 960

Qy        961 KRNTNRRPQDVKFPGGGQIVGRRGPPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSW 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 KRNTNRRPQDVKFPGGGQIVGRRGPPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSW 1020

Qy       1021 GKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGRRGPPIPKARRPEGRTWAQPGYPWPLYG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGRRGPPIPKARRPEGRTWAQPGYPWPLYG 1080

Qy       1081 NKDRRSTGKSWGKPGYPWP 1099
              |||||||||||||||||||
Db       1081 NKDRRSTGKSWGKPGYPWP 1099

RESULT 3
```

```
AAE18690
ID   AAE18690 standard; protein; 829 AA.
XX
AC   AAE18690;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Multiple epitope fusion antigen (MEFA) 12 protein.
XX
KW   Hepatitis C virus; NS3/4a antigen; multiple epitope fusion antigen;
KW   HCV infection; MEFA 12 protein.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 315
FT                   /note= "Encoded by ATG"
FT   Misc-difference 645
FT                   /note= "Encoded by GAG"
XX
PN   WO200196875-A2.
XX
PD   20-DEC-2001.
XX
PF   14-JUN-2001; 2001WO-US019369.
XX
PR   15-JUN-2000; 2000US-0212082P.
PR   02-APR-2001; 2001US-0280811P.
PR   02-APR-2001; 2001US-0280867P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Chien DY,  Arcangel P,  Tandeske L,  George-Nasciemento C,  Coit D;
PI   Medina-Selby A;
XX
DR   WPI; 2002-179522/23.
DR   N-PSDB; AAD29796.
XX
PT   Immunoassay solid support useful for detecting hepatitis C virus
PT   infection in a biological sample, comprises at least one of HCV anti-core
PT   antibody and HCV NS3/4a epitope, bound to the support.
XX
PS   Disclosure; Fig 7; 87pp; English.
XX
CC   The present invention relates to hepatitis C virus (HCV) core antigen and
CC   NS (nonstructural) 3/4a antibody combination assay that can detect both
CC   HCV antigens and antibodies present in a sample using a single solid
CC   matrix as well as immunoassay solid supports for use in the assay. The
CC   solid support is useful for detecting HCV infection in a biological
CC   sample. The present sequence is MEFA (multiple epitope fusion antigen) 12
CC   protein. This sequence is used in the exemplification of the invention
XX
SQ   Sequence 829 AA;

  Query Match             68.2%;  Score 4032;  DB 5;  Length 829;
  Best Local Similarity   69.8%;  Pred. No. 2e-275;
  Matches  791;  Conservative    1;  Mismatches    3;  Indels  338;  Gaps    7;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              ||||||||| :
Db         61 AGPHFNPLSTR---------------------------------------------------- 71

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180
                                                 ||||||||||||||||||||||
Db         72 ----------------------------------------GCNCSIYPGHITGHRMAWKLGS 93

Qy        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLITS-------- 145

Qy        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 --------TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 197

Qy        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 257

Qy        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 317

Qy        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 377

Qy        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540
              ||||||||||||||
Db        378 GYTGDFDSVIDCNTC---------------------------------------------- 392

Qy        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600

Db        393 ---------------------------------------------------------- 392

Qy        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660

Db        393 ---------------------------------------------------------- 392

Qy        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSACSGKPAIIPDREVLYREFDEMEE 720
                                                  ||||||||||||||||||||||
Db        393 ----------------------------------ACSGKPAIIPDREVLYREFDEMEE 416

Qy        721 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 476

Qy        781 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 536

Qy        841 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 596

Qy        901 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 656

Qy        961 KRNTNRRPQDVKFPGGGQIVG------RRGP-----------PIPKARRPEGRTWAQPGY 1003
              |||||||||||||||||||||      ||||           ||||||||||||||||||
Db        657 KRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVLATRKTSPIPKARRPEGRTWAQPGY 716

Qy       1004 PWPLYGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVG------RRGP- 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||      ||||
Db        717 PWPLYGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPR 776

Qy       1057 ----------PIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 1099
                        ||||||||||||||||||||||||||||||||||||||||||||
Db        777 LGVLATRKTSPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 829

RESULT 4
ADC06769
ID   ADC06769 standard; protein; 829 AA.
XX
AC   ADC06769;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Chimeric multiple epitope fusion antigen 12 protein.
XX
KW   immunoassay solid support; HCV; NS3/4a; non-structural;
```

```
KW   non-A, non-B hepatitis; NANB; multiple epitope fusion antigen 12; MEFA12;
KW   chimeric.
XX
OS   Chimeric.
OS   Synthetic.
OS   Unidentified.
OS   Hepatitis C virus.
OS   Homo sapiens.
XX
PN   US2002192639-A1.
XX
PD   19-DEC-2002.
XX
PF   14-JUN-2001; 2001US-00881239.
XX
PR   15-JUN-2000; 2000US-0212082P.
PR   02-APR-2001; 2001US-0280811P.
PR   02-APR-2001; 2001US-0280867P.
XX
PA   (CHIE/) CHIEN D Y.
PA   (ARCA/) ARCANGEL P.
PA   (TAND/) TANDESKE L.
PA   (GEOR/) GEORGE-NASCIMENTO C.
PA   (COIT/) COIT D.
PA   (MEDI/) MEDINA-SELBY A.
XX
PI   Chien DY,  Arcangel P,  Tandeske L,  George-Nascimento C,  Coit D;
PI   Medina-Selby A;
XX
DR   WPI; 2003-644609/61.
DR   N-PSDB; ADC06770.
XX
PT   Immunoassay solid support for detecting hepatitis C virus infection in
PT   biological samples, comprises a hepatitis C virus anti-core antibody and
PT   an isolated hepatitis C virus NS3/4a epitope bound HCV anti-core
PT   antibody.
XX
PS   Claim 45; Fig 7; 40pp; English.
XX
CC   The invention relates to a novel immunoassay solid support comprising at
CC   least one hepatitis C virus (HCV) anti-core antibody and at least one
CC   isolated HCV NS3/4a (non-structural protein 3/4a) epitope bound thereto.
CC   The system of the invention may be useful for detecting HCV infection in
CC   a biological sample and for treating or detecting non-A, non-B hepatitis
CC   (NANB hepatitis). The current sequence is that of the chimeric multiple
CC   epitope fusion antigen 12 (MEFA12) protein of the invention.
XX
SQ   Sequence 829 AA;

  Query Match           68.2%;  Score 4032;  DB 7;  Length 829;
  Best Local Similarity   69.8%;  Pred. No. 2e-275;
  Matches  791;  Conservative    1;  Mismatches    3;  Indels  338;  Gaps    7;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              |||||||||  :
Db         61 AGPHFNPLSTR---------------------------------------------------- 71

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180
                                                  ||||||||||||||||||||||
Db         72 ------------------------------------GCNCSIYPGHITGHRMAWKLGS 93

Qy        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLITS-------- 145

Qy        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 --------TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 197

Qy        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 257

Qy        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 317

Qy        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 377

Qy        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540
              ||||||||||||||
Db        378 GYTGDFDSVIDCNTC---------------------------------------------- 392

Qy        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600

Db        393 ---------------------------------------------------------- 392

Qy        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660

Db        393 ---------------------------------------------------------- 392

Qy        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSACSGKPAIIPDREVLYREFDEMEE 720
                                                  ||||||||||||||||||||||
Db        393 ----------------------------------ACSGKPAIIPDREVLYREFDEMEE 416

Qy        721 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 476

Qy        781 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 536

Qy        841 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 596

Qy        901 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 656

Qy        961 KRNTNRRPQDVKFPGGGQIVG------RRGP----------PIPKARRPEGRTWAQPGY 1003
              |||||||||||||||||||||      ||||          ||||||||||||||||||
Db        657 KRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVLATRKTSPIPKARRPEGRTWAQPGY 716

Qy       1004 PWPLYGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVG------RRGP- 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||      ||||
Db        717 PWPLYGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPR 776

Qy       1057 ----------PIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 1099
                         ||||||||||||||||||||||||||||||||||||||||||
Db        777 LGVLATRKTSPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 829


RESULT 5
AAW34481
ID   AAW34481 standard; protein; 1021 AA.
XX
AC   AAW34481;
XX
DT   25-MAR-2003  (revised)
DT   16-MAR-1998  (first entry)
XX
DE   HCV antigen combination pSOD/c200/core.
XX
KW   PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
KW   C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
KW   NS4.
XX
OS   Hepatitis C virus.
```

```
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .902
FT                   /note= "linker"
FT   Misc-difference 1. .154
FT                   /note= "hSOD fragment"
FT   Misc-difference 155. .159
FT                   /note= "linker"
FT   Misc-difference 160. .899
FT                   /note= "c200 (amino acids 1192-1931 of HCV polyprotein)"
FT   Misc-difference 903. .1021
FT                   /note= "c22 (amino acids 2-120 of HCV polyprotein)"
XX
PN   US5683864-A.
XX
PD   04-NOV-1997.
XX
PF   07-JUL-1992;   92US-00910760.
XX
PR   18-NOV-1987;   87US-00122714.
PR   30-DEC-1987;   87US-00139886.
PR   26-FEB-1988;   88US-00161072.
PR   06-MAY-1988;   88US-00191263.
PR   26-OCT-1988;   88US-00263584.
PR   14-NOV-1988;   88US-00271450.
PR   17-MAR-1989;   89US-00325338.
PR   20-APR-1989;   89US-00341334.
PR   21-APR-1989;   89US-00353896.
PR   18-MAY-1989;   89US-00355002.
PR   04-APR-1990;   90US-00504352.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kuo G,  Houghton M,  Choo Q;
XX
DR   WPI; 1997-548976/50.
DR   N-PSDB; AAT99982.
XX
PT   Combination of three hepatitis C virus antigens - used for detection of
PT   specific antibodies to diagnose infection.
XX
PS   Example 6; Col 59-68; 57pp; English.
XX
CC   This sequence represents a Hepatitis c virus (HCV) antigen combination of
CC   the invention. The HCV antigen combination comprises an antigen (Ag1)
CC   comprising the C domain (i.e. amino acids (aa) 1-120 of the HCV
CC   polyprotein), or its immunologically reactive fragment containing at
CC   least 8 aa. It also comprises two additional antigens from two different
CC   polyprotein domains, including at least 8 aa from the NS3, NS4, S or NS5
CC   domains of the polyprotein, corresponding, respectively, to aa 1050-1640;
CC   1640-2000; 120-400 and 2000-3011 of the HCV polyprotein. Alternatively,
CC   Ag1 contains at least 8 aa from the 1-122 or 9-177 aa regions of the HCV
CC   polyprotein. These antigen combinations are used diagnostically to detect
CC   anti-HCV antibodies, using any standard immunoassay format. These antigen
CC   combinations have a broader range of reactivity with antibodies than any
CC   antigen individually. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ   Sequence 1021 AA;

  Query Match             64.8%;  Score 3829.5;  DB 2;  Length 1021;
  Best Local Similarity   67.2%;  Pred. No. 5.2e-261;
  Matches  784;  Conservative   24;  Mismatches   79;  Indels  279;  Gaps   17;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLEFGA---------------------- 160

Qy        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240
                                                          ||||||||||
Db        161 ----------------------------------------------------VDFIPVENLE 170

Qy        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 230

Qy        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 290

Qy        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 350

Qy        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 410

Qy        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 470

Qy        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 530

Qy        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 590

Qy        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTS------------------------ 696
              ||||||||||||||||||||||||||||||||
Db        591 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVI 650

Qy        697 ----ACSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRG 748
                  ||||||||||||||||||||||||||||||||||||||||||||||||||    ||
Db        651 VGRVVLSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQ 710

Qy        749 GKPAIVPDKEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVT 803
               :  | |  :  :|: :      |        |:     :       :  |:     | :
Db        711 AE-VIAPAVQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTS 767

Qy        804 P---DKEILYE----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGL 843
              |    : :|:                ||      :    :  |: : ::|      |:
Db        768 PLTTSQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGV 827

Qy        844 LG-----------------------------------ILRRHVGPGEGAVQWMNRL 864
               |                                   ||||||||||||||||||
Db        828 AGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRL 887

Qy        865 IAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRSSRR 924
              ||||||||||||                                        | ||
Db        888 IAFASRGNHVSP----------------------------------------GNSST- 904

Qy        925 FAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNTNRRPQDVKFPGGGQIVG--- 981
                          ||        ||        :| ||||||||||||||||||
Db        905 -------------NP--------KPQ---------KKNKRNTNRRPQDVKFPGGGQIVGGVY 936

Qy        982 ---RRGP--------------------PIPKARRPEGRTWAQPGYPWPLYGNK------ 1011
                  ||||                  ||||||||||||||||||||||:
Db        937 LLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGWAG 996

Qy       1012 ---DRRSTGKSWGKPGYPWPRKTKRN 1034
                  | :  |||      ||:  ||
Db        997 WLLSPRGSRPSWGPTD---PRRRSRN 1019

RESULT 6
```

```
AAW40039
ID   AAW40039 standard; protein; 1021 AA.
XX
AC   AAW40039;
XX
DT   26-MAY-1998  (first entry)
XX
DE   Fusion protein c200/c22.
XX
KW   Hepatitis C virus C domain; HCV; immunological activity; c200/c22;
KW   NS3 domain; NS4 domain; S domain; NS5 domain; fusion protein.
XX
OS   Synthetic.
OS   Hepatitis virus.
XX
PN   US5712087-A.
XX
PD   27-JAN-1998.
XX
PF   12-MAY-1995;   95US-00440519.
XX
PR   04-APR-1990;   90US-00504352.
PR   07-JUL-1992;   92US-00910760.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kuo G,  Houghton M,  Choo Q;
XX
DR   WPI; 1998-119973/11.
DR   N-PSDB; AAV09990.
XX
PT   Immunoassays for hepatitis C virus antibodies - using combinations of
PT   antigenic fragments of HCV polyprotein.
XX
PS   Example 6; Fig 4; 59pp; English.
XX
CC   This sequence represents a fusion protein constructed from the hepatitis
CC   C virus core domain (which is situated at the carboxy terminus of the
CC   fusion protein) and a c200 construct (a fusion of the NS3 and NS3
CC   domains). This protein used in the construction of novel combinations of
CC   HCV antigens that have a broader range of immunological activity than any
CC   single HCV antigen. An example of such an antigen given in this
CC   specification comprises a first antigen containing at least 8 amino acids
CC   of the C domain of the HCV polyprotein and a second antigen comprising at
CC   least 8 amino acids of the NS3 domain, the NS4 domain, the S domain or
CC   the NS5 domain of the HCV polyprotein in the form of a fusion protein, a
CC   physical mixture or bound to a solid matrix
XX
SQ   Sequence 1021 AA;

  Query Match             64.8%;  Score 3829.5;  DB 2;  Length 1021;
  Best Local Similarity   67.2%;  Pred. No. 5.2e-261;
  Matches  784;  Conservative   24;  Mismatches   79;  Indels  279;  Gaps   17;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180
              ||||||||||||||||||||||||||||||||||   |
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLEFGA-------------------- 160

Qy        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240
                                                          ||||||||||
Db        161 ------------------------------------------------VDFIPVENLE 170

Qy        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 230

Qy        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 290

Qy        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 350

Qy        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 410

Qy        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 470

Qy        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 530

Qy        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 590

Qy        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTS---------------------- 696
              ||||||||||||||||||||||||||||||||||||
Db        591 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVI 650

Qy        697 ----ACSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRG 748
                  ||||||||||||||||||||||||||||||||||||||||||||||||||    ||
Db        651 VGRVVLSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQ 710

Qy        749 GKPAIVPDKEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVT 803
                : | |  :  :|: :       |        |:  :      :  |:    | :
Db        711 AE-VIAPAVQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTS 767

Qy        804 P---DKEILYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGL 843
              |    : :|:                 ||      :    :  |: : ::|     |:
Db        768 PLTTSQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGV 827

Qy        844 LG-------------------------------------ILRRHVGPGEGAVQWMNRL 864
              |                                     ||||||||||||||||||
Db        828 AGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRL 887

Qy        865 IAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRSSRR 924
              ||||||||||||                                          | ||
Db        888 IAFASRGNHVSP-------------------------------------------GNSST- 904

Qy        925 FAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNTNRRPQDVKFPGGGQIVG--- 981
                          ||      ||         :| ||||||||||||||||||||
Db        905 ------------NP-------KPQ--------KKNKRNTNRRPQDVKFPGGGQIVGGVY 936

Qy        982 ---RRGP--------------------PIPKARRPEGRTWAQPGYPWPLYGNK------ 1011
                 ||||                   |||||||||||||||||||||:
Db        937 LLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGWAG 996

Qy       1012 ---DRRSTGKSWGKPGYPWPRKTKRN 1034
                  | :  |||       ||:  ||
Db        997 WLLSPRGSRPSWGPTD---PRRRSRN 1019


RESULT 7
AAE22050
ID   AAE22050 standard; protein; 1021 AA.
XX
AC   AAE22050;
XX
DT   16-JUL-2002  (first entry)
XX
DE   pSOD/c200/core expression plasmid protein.
XX
KW   Hepatitis C virus; HCV; antigen; C domain; polyprotein; NS3 domain;
```

```
KW   NS4 domain; S domain; NS5 domain; pSOD/c200/core plasmid.
XX
OS   Hepatitis C virus.
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .154
FT                   /note= "hSOD"
FT   Region          155. .159
FT                   /note= "Linker region"
FT   Region          160. .899
FT                   /note= "HCV c200"
FT   Region          900. .902
FT                   /note= "Linker region"
FT   Region          903. .1021
FT                   /note= "HCV c22"
XX
PN   US6312889-B1.
XX
PD   06-NOV-2001.
XX
PF   12-MAY-1995;   95US-00440549.
XX
PR   04-APR-1990;   90US-00504352.
PR   07-JUL-1992;   92US-00910760.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Houghton M,  Choo Q,  Kuo G;
XX
DR   WPI; 2002-040268/05.
DR   N-PSDB; AAD35044.
XX
PT   Combination of hepatitis C viral (HCV) antigens, useful in improved
PT   immunoassay for detecting HCV antibodies.
XX
PS   Example 6; Fig 4; 58pp; English.
XX
CC   The invention relates to combination of hepatitis C viral (HCV) antigens
CC   that have a broader range of immunological reactivity than any single HCV
CC   antigen. The combinations consist of an antigen from the C domain of the
CC   HCV polyprotein, and at least one additional HCV antigen from either the
CC   NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
CC   the form of fusion protein, a simple physical mixture, or the individual
CC   antigens commonly bound to a solid matrix. The combinations of antigens
CC   provides broad range immunoassays for anti-HCV antibodies. The invention
CC   therefore provides a method for detecting antibodies to HCV in a mammal
CC   suspected of containing such antibodies. The present sequence is a
CC   protein encoded by pSOD/c200/core expression plasmid DNA containing HCV
CC   coding sequence
XX
SQ   Sequence 1021 AA;

  Query Match             64.8%;  Score 3829.5;  DB 5;  Length 1021;
  Best Local Similarity   67.2%;  Pred. No. 5.2e-261;
  Matches  784;  Conservative   24;  Mismatches   79;  Indels  279;  Gaps   17;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180
              |||||||||||||||||||||||||||||||||| |
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLEFGA---------------------- 160

Qy        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240
                                                          ||||||||||
Db        161 ----------------------------------------------------VDFIPVENLE 170

Qy        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 230

Qy        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 290

Qy        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 350

Qy        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 410

Qy        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 470

Qy        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 530

Qy        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 590

Qy        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTS---------------------- 696
              ||||||||||||||||||||||||||||||||||
Db        591 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVI 650

Qy        697 ----ACSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRG 748
                  ||||||||||||||||||||||||||||||||||||||||||||||||    ||
Db        651 VGRVVLSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQ 710

Qy        749 GKPAIVPDKEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVT 803
               :  | |  :  :|: :      |       |:      :       :  |:     | :
Db        711 AE-VIAPAVQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTS 767

Qy        804 P---DKEILYE----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGL 843
              |    : :|:                ||      :    :  |: : ::|     |:
Db        768 PLTTSQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGV 827

Qy        844 LG-----------------------------------ILRRHVGPGEGAVQWMNRL 864
               |                                   ||||||||||||||||||
Db        828 AGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRL 887

Qy        865 IAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRSSRR 924
              ||||||||||||                                        | ||
Db        888 IAFASRGNHVSP-------------------------------------------GNSST- 904

Qy        925 FAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNTNRRPQDVKFPGGGQIVG--- 981
                          ||       ||        :| ||||||||||||||||||
Db        905 -------------NP-------KPQ--------KKNKRNTNRRPQDVKFPGGGQIVGGVY 936

Qy        982 ---RRGP--------------------PIPKARRPEGRTWAQPGYPWPLYGNK------ 1011
                  ||||                  ||||||||||||||||||||||:
Db        937 LLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGWAG 996

Qy       1012 ---DRRSTGKSWGKPGYPWPRKTKRN 1034
                  | :  |||      ||:  ||
Db        997 WLLSPRGSRPSWGPTD---PRRRSRN 1019


RESULT 8
AAR68547
ID   AAR68547 standard; protein; 841 AA.
XX
AC   AAR68547;
XX
DT   25-MAR-2003  (revised)
```

```
DT   17-AUG-1995  (first entry)
XX
DE   HCV protease/hSOD fusion protein expression vector cflSODp600.
XX
KW   Hepatitis C virus protease/hSOD fusion protein; HCV;
KW   expression vector cflSODp600; viral infectivity inhibition.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .155
FT                   /label= SOD leader
XX
PN   US5371017-A.
XX
PD   06-DEC-1994.
XX
PF   04-APR-1991;   91US-00680296.
XX
PR   04-APR-1990;   90US-00505433.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Houghton M,  Choo Q,  Kuo G;
XX
DR   WPI; 1995-021889/03.
DR   N-PSDB; AAQ80175.
XX
PT   DNA encoding hepatitis C virus protease - used to produce large amts. of
PT   the protease and to develop prods. for inhibition of viral infectivity.
XX
PS   Claim 10; Fig 10; 69pp; English.
XX
CC   AAQ80175 (which encodes AAR68547) describes the sequence of the hepatitis
CC   C virus (HCV) protease/hSOD fusion protein E. coli expression vector,
CC   cflSODp600. Other claimed HCV protease fusion partners are yeast alpha-
CC   factor, IL-2S, ubiquitin, beta-galactosidase, beta-lactamase, horseradish
CC   peroxidase, glucose oxidase and urease. The HCV protease fusion proteins
CC   can be used in the production of Abs. They can also be used for assaying
CC   agents which inhibit protease activity, to identify compounds which
CC   inhibit viral infectivity. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 841 AA;

  Query Match             51.6%;  Score 3050.5;  DB 2;  Length 841;
  Best Local Similarity   72.6%;  Pred. No. 3.6e-206;
  Matches  615;  Conservative   10;  Mismatches   37;  Indels  185;  Gaps   11;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATNPVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHR------- 173
              ||||||||||||||||||||||||||||||||  :      :|  |:|  |
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIRR-----IGTYVY-NHLTPLRDWAHNGL 174

Qy        174 ------------------MAWKLGSAA-------------------RTTSGFVS---- 190
                                : |  :||                      | ||
Db        175 RDLAVAVEPVVFSQMETKLITWGADTAACGDIINGLPVSARRGREILLGPADGMVSKGWR 234

Qy        191 LFAP-------------------GAKQNETH------------------------VTG 205
               | ||                  |  :|:                          |
Db        235 LLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCIINGVCWTVYH 294

Qy        206 GAAART---------------------------TSGLT--------------------- 216
               ||  ||                           |  ||
Db        295 GAGTRTIASPKGPVIQMYTNVDQDLVGWPASQGTRSLTPCTCGSSDLYLVTRHADVIPVR 354

Qy        217 -------SLFSP-------GAS--------------------QNIQLIVDFIPVENLET 241
                 || ||        |:|                     : :   ||||||||||
Db        355 RRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLET 414

Qy        242 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 474

Qy        302 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 534

Qy        362 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 594

Qy        422 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 654

Qy        482 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 714

Qy        542 PGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 601
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 PGERPPGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 774

Qy        602 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 834

Qy        662 LLYRLGA 668
              |||||||
Db        835 LLYRLGA 841


RESULT 9
ABO27020
ID   ABO27020 standard; protein; 841 AA.
XX
AC   ABO27020;
XX
DT   09-SEP-2003  (first entry)
XX
DE   Hepatitis C virus (HCV) protease associated vector cflSODp600.
XX
KW   Hepatitis C virus; HCV; protease; protease inhibition; viral infection;
KW   enzyme.
XX
OS   Synthetic.
XX
PN   US2003027317-A1.
XX
PD   06-FEB-2003.
XX
PF   18-JUN-2001; 2001US-00884456.
XX
PR   04-APR-1990;   90US-00505433.
PR   04-APR-1991;   91US-00680296.
PR   06-DEC-1994;   94US-00350884.
PR   12-MAY-1995;   95US-00440548.
PR   06-SEP-1996;   96US-00709177.
PR   19-FEB-1999;   99US-00253230.
XX
PA   (HOUG/) HOUGHTON M.
PA   (CHOO/) CHOO Q.
PA   (KUOG/) KUO G.
XX
PI   Houghton M,  Choo Q,  Kuo G;
XX
DR   WPI; 2003-492037/58.
DR   N-PSDB; ACD44796.
XX
```

```
PT   New compositions comprising a hepatitis C virus (HCV) protease
PT   polynucleotide, useful for assaying pharmaceutical agents for controlling
PT   HCV, and as compounds which inhibit the protease activity and viral
PT   infectivity.
XX
PS   Disclosure; Fig 10; 41pp; English.
XX
CC   The invention describes a composition comprising a polynucleotide which
CC   encodes only the hepatitis C virus (HCV) protease or an active HCV
CC   protease analogue. The protease is useful for assaying pharmaceutical
CC   agents for controlling HCV, and as compounds which inhibit the protease
CC   activity sufficiently will also inhibit viral infectivity. This is the
CC   amino acid sequence of hepatitis C virus (HCV) protease
XX
SQ   Sequence 841 AA;

  Query Match             51.6%;  Score 3050.5;  DB 6;  Length 841;
  Best Local Similarity   72.6%;  Pred. No. 3.6e-206;
  Matches  615;  Conservative   10;  Mismatches   37;  Indels  185;  Gaps   11;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATNPVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHR------- 173
              |||||||||||||||||||||||||||||| :         :|  |:|  |
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIRR-----IGTYVY-NHLTPLRDWAHNGL 174

Qy        174 ------------------MAWKLGSAA-------------------RTTSGFVS---- 190
                               : |   :||                    | ||
Db        175 RDLAVAVEPVVFSQMETKLITWGADTAACGDIINGLPVSARRGREILLGPADGMVSKGWR 234

Qy        191 LFAP-------------------GAKQNETH-------------------------VTG 205
              | ||                  |  :|:                         |
Db        235 LLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCIINGVCWTVYH 294

Qy        206 GAAART--------------------------TSGLT---------------------- 216
              ||  ||                         |  ||
Db        295 GAGTRTIASPKGPVIQMYTNVDQDLVGWPASQGTRSLTPCTCGSSDLYLVTRHADVIPVR 354

Qy        217 -------SLFSP-------GAS--------------------QNIQLIVDFIPVENLET 241
                    || ||       |:|                    : :   ||||||||||
Db        355 RRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLET 414

Qy        242 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 474

Qy        302 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 534

Qy        362 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 421
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 594

Qy        422 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 481
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 654

Qy        482 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 541
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 714

Qy        542 PGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 601
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 PGERPPGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 774

Qy        602 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 661
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 834

Qy        662 LLYRLGA 668
              |||||||
Db        835 LLYRLGA 841


RESULT 10
ADA07875
ID   ADA07875 standard; protein; 841 AA.
XX
AC   ADA07875;
XX
DT   06-NOV-2003  (first entry)
XX
DE   HCV protease/hSOD fusion protein.
XX
KW   HCV; virucide; NS3 protease; serine protease; hSOD; superoxide dismutase;
KW   yeast a-factor; interleukin-2S; ubiquitin; beta-galactosidase;
KW   beta-lactamase; horseradish peroxidase; glucose oxidase; urease;
KW   HCV infection; cf1SODp600; human.
XX
OS   Chimeric.
OS   Hepatitis C virus.
OS   Homo sapiens.
XX
PN   US2003064499-A1.
XX
PD   03-APR-2003.
XX
PF   18-JUN-2001; 2001US-00884455.
XX
PR   04-APR-1990;   90US-00505433.
PR   04-APR-1991;   91US-00680296.
PR   06-DEC-1994;   94US-00350884.
PR   12-MAY-1995;   95US-00440548.
PR   06-SEP-1996;   96US-00709177.
PR   18-FEB-1999;   99US-00253675.
XX
PA   (HCUG/) HOUGHTON M.
PA   (CHOO/) CHOO Q.
PA   (KUOG/) KUO G.
XX
PI   Houghton M,  Choo Q,  Kuo G;
XX
DR   WPI; 2003-540789/51.
DR   N-PSDB; ADA07874.
XX
PT   A composition for assaying and designing antiviral agents specific for
PT   Hepatitis C virus (HCV) comprises a purified proteolytic polypeptide from
PT   HCV or a polynucleotide which encodes HCV protease.
XX
PS   Disclosure; Fig 10; 40pp; English.
XX
CC   The invention relates to a composition comprising a purified proteolytic
CC   polypeptide derived from Hepatitis C virus (HCV) or a polynucleotide
CC   which encodes only the HCV protease or an active HCV protease analogue,
CC   or which encodes a fusion protein comprising HCV protease or HCV protease
CC   analogue, and a fusion partner. Also included are a fusion protein
CC   comprising a fusion partner fused to a proteolytic polypeptide derived
CC   from HCV, a method for assaying compounds for activity against HCV
CC   (comprising providing an active HCV protease, contacting the protease
CC   with a compound capable of inhibiting serine protease activity and
CC   measuring inhibition of the proteolytic activity of the HCV protease) and
CC   an expression vector for producing HCV protease or HCV protease analogues
CC   in a host cell (comprising a polynucleotide encoding HCV protease or an
CC   HCV protease analogue, transcriptional and translational regulatory
CC   sequences functional in the host cell operably linked to the HCV protease
CC   -encoding polynucleotide and a selectable marker). The fusion partner is
CC   selected from hSOD (human superoxide dismutase), yeast a-factor,
CC   interleukin (IL)-2S, ubiquitin, beta-galactosidase, beta-lactamase,
CC   horseradish peroxidase, glucose oxidase and urease. The composition is
```

```
CC   useful in assaying and designing antiviral agents specific for HCV. The
CC   method is used in identifying antiviral agents effective for treating
CC   HCV. The present sequence is an HCV protease/hSOD fusion protein.
XX
SQ   Sequence 841 AA;

  Query Match             51.6%;  Score 3050.5;  DB 7;  Length 841;
  Best Local Similarity   72.6%;  Pred. No. 3.6e-206;
  Matches  615;  Conservative   10;  Mismatches   37;  Indels  185;  Gaps   11;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATNPVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHR------- 173
              |||||||||||||||||||||||||||||| :         :|  |:|  |
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIRR-----IGTYVY-NHLTPLRDWAHNGL 174

Qy        174 ------------------MAWKLGSAA--------------------RTTSGFVS---- 190
                               : |   :||                    | ||
Db        175 RDLAVAVEPVVFSQMETKLITWGADTAACGDIINGLPVSARRGREILLGPADGMVSKGWR 234

Qy        191 LFAP--------------------GAKQNETH-------------------------VTG 205
              | ||                   |  :|:                         |
Db        235 LLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCIINGVCWTVYH 294

Qy        206 GAAART---------------------------TSGLT---------------------- 216
              ||  ||                          |  ||
Db        295 GAGTRTIASPKGPVIQMYTNVDQDLVGWPASQGTRSLTPCTCGSSDLYLVTRHADVIPVR 354

Qy        217 -------SLFSP-------GAS--------------------QNIQLIVDFIPVENLET 241
                    || ||       |:|                    : :   ||||||||||
Db        355 RRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLET 414

Qy        242 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 474

Qy        302 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 534

Qy        362 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 421
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 594

Qy        422 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 481
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 654

Qy        482 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 541
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 714

Qy        542 PGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 601
              |||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 PGERPPGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 774

Qy        602 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 661
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 834

Qy        662 LLYRLGA 668
              |||||||
Db        835 LLYRLGA 841


RESULT 11
AAW01701
```

```
ID   AAW01701 standard; protein; 841 AA.
XX
AC   AAW01701;
XX
DT   17-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   03-APR-1997  (first entry)
XX
DE   hSOD-HCV fusion protein.
XX
KW   HCV; NS3; non-structural domain 3; protease; polyprotein; inhibitor;
KW   screen; processing; infection; treatment; probe; hepatitis C virus.
XX
OS   Hepatitis C virus; Virus.
OS   Homo sapiens.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Protein         156. .841
FT                   /label= HCV_protease
XX
PN   US5585258-A.
XX
PD   17-DEC-1996.
XX
PF   06-DEC-1994;   94US-00350884.
XX
PR   04-APR-1990;   90US-00505433.
PR   04-APR-1991;   91US-00680296.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Choo Q,  Kuo G,  Houghton M;
XX
DR   WPI; 1997-051175/05.
DR   N-PSDB; AAT59261.
XX
PT   Compsn. contg. hepatitis C virus NS3 domain protease and related fusion
PT   proteins - useful for screening specific inhibitors, potential antiviral
PT   agents, prepn. of antibodies and for cleaving specific poly:peptide(s).
XX
PS   Example 4; Col 77-84; 68pp; English.
XX
CC   Compsns. comprising the hepatitis C virus (HCV) NS3 domain protease or
CC   its active truncation analogues are claimed. Also new are fusion proteins
CC   comprising the protease (or analogues) and, e.g. human superoxide (SOD)
CC   or ubiquitin. The protease is essential for polyprotein processing, and
CC   thus infectivity, in HCV. The compsns. are used to screen for specific
CC   inhibitors (possibly useful as antiviral agents), to generate specific
CC   antibodies and to cleave specific polypeptides. HCV cDNA clones (AAT59250
CC   - 56 encoding AAW01686-92 resp.) were isolated from HCV genomic library
CC   using probes AAT59244-49. The clones were used in the preparation of full
CC   -length SOD-protease fusion proteins. The present sequence is encoded by
CC   vector cf1SODp600 which contains a full-length HCV protease coding
CC   sequence fused to a functional hSOD leader. The resulting vector encodes
CC   amino acids 1-151 of hSOD, and amino acids 946-1630 of HCV (corresponding
CC   to 1-686 of AAW01693). (Updated on 25-MAR-2003 to correct PF field.)
CC   (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 841 AA;

  Query Match             51.5%;  Score 3047.5;  DB 2;  Length 841;
  Best Local Similarity   72.6%;  Pred. No. 5.9e-206;
  Matches  615;  Conservative    9;  Mismatches   38;  Indels  185;  Gaps   11;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATNPVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
```

```
Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHR------- 173
              ||||||||||||||||||||||||||||||||||||          :|  |:|  |
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIRIR-----GTYVY-NHLTPLRDWAHNGL 174

Qy        174 ------------------MAWKLGSAA--------------------RTTSGFVS---- 190
                               : |   :||                    | ||
Db        175 RDLAVAVEPVVFSQMETKLITWGADTAACGDIINGLPVSARRGREILLGPADGMVSKGWR 234

Qy        191 LFAP--------------------GAKQNETH-------------------------VTG 205
              | ||                   | :|:                         |
Db        235 LLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCIINGVCWTVYH 294

Qy        206 GAAART---------------------------TSGLT---------------------- 216
              ||  ||                          |  ||
Db        295 GAGTRTIASPKGPVIQMYTNVDQDLVGWPASQGTRSLTPCTCGSSDLYLVTRHADVIPVR 354

Qy        217 -------SLFSP-------GAS--------------------QNIQLIVDFIPVENLET 241
                    || ||       |:|                    : :   ||||||||||
Db        355 RRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLET 414

Qy        242 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 474

Qy        302 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 534

Qy        362 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 594

Qy        422 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 654

Qy        482 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 714

Qy        542 PGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 601
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 PGERPPGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 774

Qy        602 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 834

Qy        662 LLYRLGA 668
              |||||||
Db        835 LLYRLGA 841
```

RESULT 12
AAW46397
ID AAW46397 standard; protein; 841 AA.
XX
AC AAW46397;
XX
DT 27-AUG-2003 (revised)
DT 07-MAY-1998 (first entry)
XX
DE Amino acid sequence of the vector cf1SODp600.
XX
KW Protease; HCV; NS3 domain; human superoxide dismutase; fusion protein;
KW assay; activity; anti-HCV.
XX
OS Synthetic.
OS Hepatitis C virus.
OS Homo sapiens.
XX
PN US5712145-A.
XX
PD 27-JAN-1998.
XX
PF 06-SEP-1996; 96US-00709173.
XX
PR 04-APR-1990; 90US-00505433.
PR 04-APR-1991; 91US-00680296.
PR 06-DEC-1994; 94US-00350884.
PR 12-MAY-1995; 95US-00440548.
XX
PA (CHIR) CHIRON CORP.
XX
PI Choo Q, Kuo G, Houghton M;
XX
DR WPI; 1998-119986/11.
DR N-PSDB; AAV04993.
XX
PT Recombinant hepatitis C virus protease - useful in screening drugs for
PT activity against hepatitis C virus.
XX
PS Disclosure; Fig 10A-G; 68pp; English.
XX
CC The present sequence represents the amino acid sequence of the vector
CC cf1SODp600. This vector contains a full length Hepatitis C virus (HCV)
CC protease coding sequence fused to a functional human superoxide dismutase
CC leader. The vector was used to express the protease fusion protein in
CC Escherichia coli. The HCV protease is believed to cleave itself from the
CC genomic polyprotein. In the absence of protease activity, the HCV
CC polyprotein should remain in its unprocessed form, and thus render the
CC virus non-infectious. Inhibitors of protease activity should also inhibit
CC viral infectivity. The protease can therefore be used for assaying
CC compounds for activity against HCV. (Updated on 27-AUG-2003 to correct OS
CC field.)
XX
SQ Sequence 841 AA;

Query Match 51.5%; Score 3047.5; DB 2; Length 841;
Best Local Similarity 72.6%; Pred. No. 5.9e-206;
Matches 615; Conservative 9; Mismatches 38; Indels 185; Gaps 11;

```
Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATNPVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHR------- 173
              |||||||||||||||||||||||||||||||||||          :|  |:|  |
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIRIR-----GTYVY-NHLTPLRDWAHNGL 174

Qy        174 ------------------MAWKLGSAA--------------------RTTSGFVS---- 190
                               : |   :||                    | ||
Db        175 RDLAVAVEPVVFSQMETKLITWGADTAACGDIINGLPVSARRGREILLGPADGMVSKGWR 234

Qy        191 LFAP--------------------GAKQNETH-------------------------VTG 205
              | ||                   | :|:                         |
Db        235 LLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCIINGVCWTVYH 294

Qy        206 GAAART---------------------------TSGLT---------------------- 216
              ||  ||                          |  ||
Db        295 GAGTRTIASPKGPVIQMYTNVDQDLVGWPASQGTRSLTPCTCGSSDLYLVTRHADVIPVR 354

Qy        217 -------SLFSP-------GAS--------------------QNIQLIVDFIPVENLET 241
                    || ||       |:|                    : :   ||||||||||
Db        355 RRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLET 414

Qy        242 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 474

Qy        302 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        475 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 534

Qy        362 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 594

Qy        422 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 654

Qy        482 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 714

Qy        542 PGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 601
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 PGERPPGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 774

Qy        602 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 834

Qy        662 LLYRLGA 668
              |||||||
Db        835 LLYRLGA 841
```

```
RESULT 13
AAW97609
ID   AAW97609 standard; protein; 841 AA.
XX
AC   AAW97609;
XX
DT   26-MAY-1999  (first entry)
XX
DE   Amino acid seqeunce of vector cflSODp600.
XX
KW   HCV NS3 protease; truncation analog; HCV control; protease activity;
KW   viral infectivity; inactive non-cleaving protease.
XX
OS   Synthetic.
OS   Hepatitis C virus.
XX
PN   US5885799-A.
XX
PD   23-MAR-1999.
XX
PF   06-SEP-1996;   96US-00709177.
XX
PR   04-APR-1990;   90US-00505433.
PR   04-APR-1991;   91US-00680296.
PR   06-DEC-1994;   94US-00350884.
PR   12-MAY-1995;   95US-00440548.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Choo Q,  Kuo G,  Houghton M;
XX
DR   WPI; 1999-228536/19.
DR   N-PSDB; AAX26398.
XX
PT   Preparation of new Hepatitis C Virus NS3 protease - useful for screening
PT   for compounds which inhibit HCV infectivity.
XX
PS   Example 3; Fig 10; 71pp; English.
XX
CC   The specification describes a method for making a purified Hepatitis C
CC   virus (HCV) NS3 protease or active truncation analog. If the HCV protease
CC   N-terminal cleavage signal is excluded (so that self-cleavage is
CC   prevented), the HCV protease remains in its unprocessed form, and renders
CC   the virus noninfectious. The protease is therefore useful for assaying
CC   pharmaceutical agents for control of HCV, as compounds which inhibit
CC   protease activity sufficiently will also inhibit viral infectivity. An
CC   inactive non-cleaving protease can be used to screen for inhibitors.
CC   Recombinant expression systems can be utilised to prepare recombinant HCV
CC   which can be used to produce monoclonal antibodies. The present sequence
CC   was created in the course of the invention
XX
SQ   Sequence 841 AA;
```

```
  Query Match             51.5%;  Score 3047.5;  DB 2;  Length 841;
  Best Local Similarity   72.6%;  Pred. No. 5.9e-206;
  Matches  615;  Conservative    9;  Mismatches   38;  Indels  185;  Gaps   11;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATNPVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHR------- 173
              |||||||||||||||||||||||||||||||||||||||||||||:|  |:|  |
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIRIR-----GTYVY-NHLTPLRDWAHNGL 174

Qy        174 -------------------MAWKLGSAA--------------------RTTSGFVS---- 190
                                  : |   :||                    | ||
Db        175 RDLAVAVEPVVFSQMETKLITWGADTAACGDIINGLPVSARRGREILLGPADGMVSKGWR 234

Qy        191 LFAP--------------------GAKQNETH-------------------------VTG 205
              | ||                    |  :|:                         |
Db        235 LLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCIINGVCWTVYH 294

Qy        206 GAAART---------------------------TSGLT---------------------- 216
              ||  ||                          |  ||
Db        295 GAGTRTIASPKGPVIQMYTNVDQDLVGWPASQGTRSLTPCTCGSSDLYLVTRHADVIPVR 354

Qy        217 -------SLFSP-------GAS--------------------QNIQLIVDFIPVENLET 241
                    || ||       |:|                    : :   ||||||||||
Db        355 RRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLET 414

Qy        242 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLG 474

Qy        302 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 FGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDA 534

Qy        362 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLE 594

Qy        422 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 VIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTG 654

Qy        482 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 YTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVA 714

Qy        542 PGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 601
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 PGERPPGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFT 774

Qy        602 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTP 834

Qy        662 LLYRLGA 668
              |||||||
Db        835 LLYRLGA 841
```

RESULT 14

```
AAR14349
ID   AAR14349 standard; protein; 840 AA.
XX
AC   AAR14349;
XX
DT   16-JAN-1992  (first entry)
XX
DE   HCV protease::hSOD leader fusion encoded by cf1SODp600.
XX
KW   Hepatitis C virus; HCV; human superoxide dismutase; SOD.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .156
FT                   /label= hSOD leader
FT   Protein         157
FT                   /label= truncated HCV protease
XX
PN   WO9115596-A.
XX
PD   17-OCT-1991.
XX
PF   04-APR-1990;   90US-00505434.
XX
PR   04-APR-1990;   90US-00505434.
XX
PA   (PROT-) PROTOS INC.
XX
PI   Rosenberg S;
XX
DR   WPI; 1991-325236/44.
DR   N-PSDB; AAQ14358.
XX
PT   Method for assaying pharmaceutical cpds. - for determining anti-Hepatitis
PT   C Virus activity, using binding affinity.
XX
PS   Example 4; Fig 10; 68pp; English.
XX
CC   The vector cf1SODp600 contains a full-length HCV protease coding sequence
CC   fused to a functional hSOD leader. The truncated protease analogue
CC   expressed by the vector is proteolitically inactive and can be used to
CC   assay a wide range of pharmaceutical agents for controlling HCV. Those
CC   agents which inhibit the protease activity sufficiently will also inhibit
CC   viral infectivity. See also AAR14350-R14356
XX
SQ   Sequence 840 AA;

  Query Match             51.5%;  Score 3042.5;  DB 2;  Length 840;
  Best Local Similarity   72.6%;  Pred. No. 1.3e-205;
  Matches  614;  Conservative    9;  Mismatches   38;  Indels  185;  Gaps   11;

Qy          2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTSA 61
              ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATNPVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTSP 60

Qy         62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVVH 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVVH 120

Qy        122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHR-------- 173
              ||||||||||||||||||||||||||||||           :|   |:|   |
Db        121 EKADDLGKGGNEESTKTGNAGSRLACGVIGIRIR-----GTYVY-NHLTPLRDWAHNGLR 174

Qy        174 -----------------MAWKLGSAA-------------------RTTSGFVS----L 191
                               :  |   :||                   | ||     |
Db        175 DLAVAVEPVVFSQMETKLITWGADTAACGDIINGLPVSARRGREILLGPADGMVSKGWRL 234

Qy        192 FAP-------------------GAKQNETH-----------------------VTGG 206
               ||                   |  :|:                        | |
Db        235 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCIINGVCWTVYHG 294

Qy        207 AAART--------------------------TSGLT---------------------- 216
                |  ||                                 |  ||
Db        295 AGTRTIASPKGPVIQMYTNVDQDLVGWPASQGTRSLTPCTCGSSDLYLVTRHADVIPVRR 354

Qy        217 ------SLFSP-------GAS------------------QNIQLIVDFIPVENLETT 242
                    || ||       |:|                  : :    ||||||||||
Db        355 RGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETT 414

Qy        243 MRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGF 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 MRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGF 474

Qy        303 GAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDAT 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDAT 534

Qy        363 SILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEV 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 SILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEV 594

Qy        423 IKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGY 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 IKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGY 654

Qy        483 TGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAP 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 TGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAP 714

Qy        543 GERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTG 602
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 GERPPGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTG 774

Qy        603 LTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPL 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 LTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPL 834

Qy        663 LYRLGA 668
              ||||||
Db        835 LYRLGA 840


RESULT 15
AAP90164
ID   AAP90164 standard; protein; 2261 AA.
XX
AC   AAP90164;
XX
DT   25-MAR-2003  (revised)
DT   01-NOV-1989  (first entry)
XX
DE   Peptide encoded by composite hepatitis C virus cDNA.
XX
KW   Hepatitis C virus; clone 12f; clone 15e; probe; vaccine.
XX
OS   Pan troglodytes.
XX
PN   GB2212511-A.
XX
PD   26-JUL-1989.
XX
PF   18-NOV-1988;   88GB-00027024.
XX
PR   18-NOV-1987;   87US-00122714.
PR   30-DEC-1987;   87US-00139886.
PR   26-FEB-1988;   88US-00161072.
PR   26-OCT-1988;   88US-00263584.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Houghton M,  Choo QL,  Kuo G;
XX
DR   WPI; 1989-215054/30.
DR   N-PSDB; AAN90331.
XX
```

PT Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
PS Disclosure; Fig 32; 30pp; English.
XX
CC The sequence is the peptide encoded by the composite hepatitis C virus
CC (HCV) cDNA of AAN90331. The polypeptides are used to diagnose HCV-induced
CC NANBH, to raise antibodies for immunoassay or treatment, or to produce
CC vaccines. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 2261 AA;

Query Match 49.2%; Score 2909.5; DB 1; Length 2261;
Best Local Similarity 53.3%; Pred. No. 1.3e-195;
Matches 622; Conservative 33; Mismatches 92; Indels 419; Gaps 16;

```
Qy       224 SQNIQLIVDFIPVENLETTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYA 283
Db       561 TRGVAKAVDFIPVENLETTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYA 620

Qy       284 AQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGC 343
Db       621 AQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGC 680

Qy       344 SGGAYDIIICDECHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEV 403
Db       681 SGGAYDIIICDECHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEV 740

Qy       404 ALSTTGEIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSV 463
Db       741 ALSTTGEIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSV 800

Qy       464 IPTSGDVVVVATDALMTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRT 523
Db       801 IPTSGDVVVVATDALMTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRT 860

Qy       524 QRRGRTGRGKPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNT 583
Db       861 QRRGRTGRGKPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNT 920

Qy       584 PGLPVCQDHLEFWEGVFTGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWD 643
Db       921 PGLPVCQDHLEFWEGVFTGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWD 980

Qy       644 QMWKCLIRLKPTLHGPTPLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTS------- 696
Db       981 QMWKCLIRLKPTLHGPTPLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGG 1040

Qy       697 --------------------ACSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLA 735
Db      1041 VLAALAAYCLSTGCVVIVGRVVLSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLA 1100

Qy       736 EQFKQKALGL----SRGGKPAIVPDKEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQ 786
Db      1101 EQFKQKALGLLQTASRQAE-VIAPAVQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG- 1158

Qy       787 FKEKVLGLIDNDQVVVTP---DKEILYE-----------------AFDEMEECASKAALI 826
Db      1159 -NPAIASLMAFTAAVTSPLTTSQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSV 1217

Qy       827 EEGQRMAEMLKSKIQGLLG-----------------------------------------IL 847
Db      1218 GLGKVLIDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAIL 1277

Qy       848 RRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPS------------------------- 882
Db      1278 RRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTAILSSLTVTQLLRRLHQ 1337

Qy       883 ---------------------------------------------------------- 882
Db      1338 WISSECTTPCSGSWLRDIWDWICEVLSDFKTWLKAKLMPQLPGIPFVSCQRGYKGVWRVD 1397

Qy       883 ---------------------------------------------------------- 882
Db      1398 GIMHTRCHCGAEITGHVKNGTMRIVGPRTCRNMWSGTFPINAYTTGPCTPLPAPNYTFAL 1457

Qy       883 ----------------------------------------------------RSRRFA-------- 888
Db      1458 WRVSAEEYVEIRQVGDFHYVTGMTTDNLKCPCQVPSPEFFTELDGVRLHRFAPPCKPLLR 1517

Qy       889 ---------------QALPVWARPDY-------------------------NPP------ 902
Db      1518 EEVSFRVGLHEYPVGSQLPCEPEPDVAVLTSMLTDPSHITAEAAGRRLARGSPPSVASSS 1577

Qy       903 ------------------------LVET---WKK------------------PDYEPPVVH 918
Db      1578 ASQLSAPSLKATCTANHDSPDAELIEANLLWRQEMGGNITRVESENKVVILDSFDPLVAE 1637

Qy       919 G------------RSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNTNR 966
Db      1638 EDEREISVPAEILRKSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVH---------- 1687

Qy       967 RPQDVKFPGGGQIVGRRGPPIPKARR 992
Db      1688 ---------GCPLPPPKSPPVPPPRK 1704
```

Search completed: June 21, 2004, 10:30:26
Job time : 73.5868 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:18:09 ; Search time 49.4736 Seconds
(without alignments)
4734.482 Million cell updates/sec

Title: US-10-658-782-4
Perfect score: 4455
Sequence: 1 MATKAVCVLKGDGPVQGIIN.....GNKDRRSTGKSWGKPGYPWP 829

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4455	100.0	829	5	AAE18690	Aae18690 Multiple
2	4455	100.0	829	7	ADC06769	Adc06769 Chimeric
3	4032	90.5	1099	5	AAU76378	Aau76378 HCV multi
4	4032	90.5	1099	6	ABG72262	Abg72262 HCV multi
5	2222	49.9	1021	2	AAW34481	Aaw34481 HCV antig
6	2222	49.9	1021	2	AAW40039	Aaw40039 Fusion pr
7	2222	49.9	1021	5	AAE22050	Aae22050 pSOD/c200
8	1627.5	36.5	1766	1	AAP92041	Aap92041 Hepatitis
9	1624.5	36.5	2261	1	AAP90164	Aap90164 Peptide e
10	1624.5	36.5	2436	1	AAP92050	Aap92050 Sequence
11	1624.5	36.5	2436	1	AAP90288	Aap90288 Peptide e
12	1624.5	36.5	2772	3	AAB18540	Aab18540 Protein e
13	1624.5	36.5	2955	2	AAY14975	Aay14975 Amino aci
14	1624.5	36.5	2955	3	AAB18541	Aab18541 Polyprote
15	1624.5	36.5	3011	2	AAR21519	Aar21519 Compiled
16	1624.5	36.5	3011	2	AAR90931	Aar90931 Hepatitis
17	1624.5	36.5	3011	2	AAW34480	Aaw34480 HCV polyp
18	1624.5	36.5	3011	2	AAW40038	Aaw40038 HCV polyp
19	1624.5	36.5	3011	5	AAE22049	Aae22049 Hepatitis
20	1623	36.4	781	2	AAR22208	Aar22208 Sequence
21	1623	36.4	781	2	AAR21565	Aar21565 HCV CKS-3
22	1623	36.4	781	2	AAR33632	Aar33632 HCV CKS-3
23	1623	36.4	781	2	AAR33594	Aar33594 HCV CKS-3
24	1623	36.4	781	2	AAR33574	Aar33574 HCV CKS-3
25	1623	36.4	781	2	AAR52690	Aar52690 HCV CKS-3
26	1623	36.4	781	4	AAB51372	Aab51372 HCV recom
27	1622.5	36.4	3011	5	AAU84597	Aau84597 HCV polyp
28	1621.5	36.4	2301	1	AAP92047	Aap92047 Sequence
29	1621.5	36.4	2772	2	AAR08123	Aar08123 Hepatitis
30	1619.5	36.4	2894	2	AAR24440	Aar24440 Composite
31	1616.5	36.3	2816	2	AAR34009	Aar34009 HCV-1 pol
32	1615.5	36.3	2435	2	AAR25135	Aar25135 HCV polyp
33	1614.5	36.2	2894	2	AAR70230	Aar70230 Composite
34	1614.5	36.2	3011	2	AAR31621	Aar31621 Hepatitis
35	1608.5	36.1	2436	2	AAR28582	Aar28582 HCV amino
36	1605.5	36.0	1786	1	AAP90158	Aap90158 Protein s
37	1600.5	35.9	2955	2	AAR08124	Aar08124 Hepatitis
38	1597.5	35.9	2202	6	AAO26783	Aao26783 Protein d
39	1597.5	35.9	2631	6	AAO26785	Aao26785 Protein d
40	1597.5	35.9	2984	4	AAE00442	Aae00442 Hepatitis
41	1597.5	35.9	3011	2	AAR40120	Aar40120 HCV genom
42	1597.5	35.9	3011	2	AAW77397	Aaw77397 Hepatitis
43	1597.5	35.9	3011	2	AAW77398	Aaw77398 Hepatitis
44	1597.5	35.9	3011	2	AAW98021	Aaw98021 Infectiou
45	1597.5	35.9	3011	2	AAW98020	Aaw98020 Infectiou

ALIGNMENTS

RESULT 1
AAE18690
ID AAE18690 standard; protein; 829 AA.
XX
AC AAE18690;
XX
DT 17-MAY-2002 (first entry)
XX
DE Multiple epitope fusion antigen (MEFA) 12 protein.
XX
KW Hepatitis C virus; NS3/4a antigen; multiple epitope fusion antigen;
KW HCV infection; MEFA 12 protein.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 315
FT /note= "Encoded by ATG"
FT Misc-difference 645
FT /note= "Encoded by GAG"
XX
PN WO200196875-A2.
XX
PD 20-DEC-2001.
XX
PF 14-JUN-2001; 2001WO-US019369.
XX
PR 15-JUN-2000; 2000US-0212082P.
PR 02-APR-2001; 2001US-0280811P.
PR 02-APR-2001; 2001US-0280867P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Chien DY, Arcangel P, Tandeske L, George-Nasciemento C, Coit D;
PI Medina-Selby A;
XX
DR WPI; 2002-179522/23.
DR N-PSDB; AAD29796.
XX
PT Immunoassay solid support useful for detecting hepatitis C virus
PT infection in a biological sample, comprises at least one of HCV anti-core
PT antibody and HCV NS3/4a epitope, bound to the support.
XX
PS Disclosure; Fig 7; 87pp; English.
XX
CC The present invention relates to hepatitis C virus (HCV) core antigen and
CC NS (nonstructural) 3/4a antibody combination assay that can detect both
CC HCV antigens and antibodies present in a sample using a single solid

```
CC   matrix as well as immunoassay solid supports for use in the assay. The
CC   solid support is useful for detecting HCV infection in a biological
CC   sample. The present sequence is MEFA (multiple epitope fusion antigen) 12
CC   protein. This sequence is used in the exemplification of the invention
XX
SQ   Sequence 829 AA;

  Query Match             100.0%;  Score 4455;  DB 5;  Length 829;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-310;
  Matches  829;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSTRGCNCSIYPGHITGHRMAWKLGSAARTTSGFVSLFAPGAKQNETHVTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGPHFNPLSTRGCNCSIYPGHITGHRMAWKLGSAARTTSGFVSLFAPGAKQNETHVTGGA 120

Qy        121 AARTTSGLTSLFSPGASQNIQLITSTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AARTTSGLTSLFSPGASQNIQLITSTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYA 180

Qy        181 AQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGC 240

Qy        241 SGGAYDIIICDECHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SGGAYDIIICDECHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEV 300

Qy        301 ALSTTGEIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALSTTGEIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSV 360

Qy        361 IPTSGDVVVVATDALMTGYTGDFDSVIDCNTCACSGKPAIIPDREVLYREFDEMEECSQH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IPTSGDVVVVATDALMTGYTGDFDSVIDCNTCACSGKPAIIPDREVLYREFDEMEECSQH 420

Qy        421 LPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQAQVIA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQAQVIA 480

Qy        481 HQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKIQGLL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 HQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKIQGLL 540

Qy        541 GILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYNPPLV 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYNPPLV 600

Qy        601 ETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNT 660

Qy        661 NRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVLATRKTSPIPKARRPEGRTWAQPGYPWPL 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVLATRKTSPIPKARRPEGRTWAQPGYPWPL 720

Qy        721 YGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVL 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 YGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVL 780

Qy        781 ATRKTSPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 829
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATRKTSPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 829


RESULT 2
ADC06769
ID   ADC06769 standard; protein; 829 AA.
XX
AC   ADC06769;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Chimeric multiple epitope fusion antigen 12 protein.
XX
KW   immunoassay solid support; HCV; NS3/4a; non-structural;
KW   non-A, non-B hepatitis; NANB; multiple epitope fusion antigen 12; MEFA12;
KW   chimeric.
XX
OS   Chimeric.
OS   Synthetic.
OS   Unidentified.
OS   Hepatitis C virus.
OS   Homo sapiens.
XX
PN   US2002192639-A1.
XX
PD   19-DEC-2002.
XX
PF   14-JUN-2001; 2001US-00881239.
XX
PR   15-JUN-2000; 2000US-0212082P.
PR   02-APR-2001; 2001US-0280811P.
PR   02-APR-2001; 2001US-0280867P.
XX
PA   (CHIE/) CHIEN D Y.
PA   (ARCA/) ARCANGEL P.
PA   (TAND/) TANDESKE L.
PA   (GEOR/) GEORGE-NASCIMENTO C.
PA   (COIT/) COIT D.
PA   (MEDI/) MEDINA-SELBY A.
XX
PI   Chien DY,  Arcangel P,  Tandeske L,  George-Nascimento C,  Coit D;
PI   Medina-Selby A;
XX
DR   WPI; 2003-644609/61.
DR   N-PSDB; ADC06770.
XX
PT   Immunoassay solid support for detecting hepatitis C virus infection in
PT   biological samples, comprises a hepatitis C virus anti-core antibody and
PT   an isolated hepatitis C virus NS3/4a epitope bound HCV anti-core
PT   antibody.
XX
PS   Claim 45; Fig 7; 40pp; English.
XX
CC   The invention relates to a novel immunoassay solid support comprising at
CC   least one hepatitis C virus (HCV) anti-core antibody and at least one
CC   isolated HCV NS3/4a (non-structural protein 3/4a) epitope bound thereto.
CC   The system of the invention may be useful for detecting HCV infection in
CC   a biological sample and for treating or detecting non-A, non-B hepatitis
CC   (NANB hepatitis). The current sequence is that of the chimeric multiple
CC   epitope fusion antigen 12 (MEFA12) protein of the invention.
XX
SQ   Sequence 829 AA;

  Query Match             100.0%;  Score 4455;  DB 7;  Length 829;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-310;
  Matches  829;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSTRGCNCSIYPGHITGHRMAWKLGSAARTTSGFVSLFAPGAKQNETHVTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGPHFNPLSTRGCNCSIYPGHITGHRMAWKLGSAARTTSGFVSLFAPGAKQNETHVTGGA 120

Qy        121 AARTTSGLTSLFSPGASQNIQLITSTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AARTTSGLTSLFSPGASQNIQLITSTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYA 180

Qy        181 AQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 AQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGC 240

Qy        241 SGGAYDIIICDECHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEV 300

Db        241 SGGAYDIIICDECHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEV 300

Qy        301 ALSTTGEIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSV 360

Db        301 ALSTTGEIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSV 360

Qy        361 IPTSGDVVVVATDALMTGYTGDFDSVIDCNTCACSGKPAIIPDREVLYREFDEMEECSQH 420

Db        361 IPTSGDVVVVATDALMTGYTGDFDSVIDCNTCACSGKPAIIPDREVLYREFDEMEECSQH 420

Qy        421 LPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQAQVIA 480

Db        421 LPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQAQVIA 480

Qy        481 HQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKIQGLL 540

Db        481 HQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKIQGLL 540

Qy        541 GILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYNPPLV 600

Db        541 GILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYNPPLV 600

Qy        601 ETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNT 660

Db        601 ETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNT 660

Qy        661 NRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVLATRKTSPIPKARRPEGRTWAQPGYPWPL 720

Db        661 NRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVLATRKTSPIPKARRPEGRTWAQPGYPWPL 720

Qy        721 YGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVL 780

Db        721 YGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVL 780

Qy        781 ATRKTSPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 829

Db        781 ATRKTSPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 829
```

```
RESULT 3
AAU76378
ID   AAU76378 standard; protein; 1099 AA.
XX
AC   AAU76378;
XX
DT   08-MAY-2002  (first entry)
XX
DE   HCV multiple epitope fusion antigen (MEFA) 7.1 protein sequence.
XX
KW   Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion;
KW   immunoassay solid support; multiple epitope fusion antigen; MEFA;
KW   non-structural protein.
XX
OS   Hepatitis C virus.
OS   Synthetic.
XX
PN   WO200196870-A2.
XX
PD   20-DEC-2001.
XX
PF   14-JUN-2001; 2001WO-US019156.
XX
PR   15-JUN-2000; 2000US-0212082P.
PR   02-APR-2001; 2001US-0280811P.
PR   02-APR-2001; 2001US-0280867P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Chien DY,  Arcangel P,  Tandeske L,  George-Nasciemento C,  Coit D;
PI   Medina-Selby A;
XX
DR   WPI; 2002-090228/12.
DR   N-PSDB; ABK15345.
XX
PT   Immunoassay solid support, useful for detecting hepatitis C virus
PT   infection in biological sample, comprises HCV NS3/4a conformational
PT   epitope and multiple epitope fusion antigen bound to the support.
XX
PS   Claim 5; Fig 5; 92pp; English.
XX
CC   The present invention relates to a new immunoassay solid support
CC   consisting essentially of at least one hepatitis C virus (HCV) NS3/4a
CC   conformational epitope and a multiple epitope fusion antigen (MEFA),
CC   bound to the support. The NS3/4a conformational epitope and/or MEFA
CC   reacts specifically with anti-HCV antibodies present in a biological
CC   sample from an HCV-infected individual. The immunoassay of the invention
CC   is useful for detecting hepatitis C virus infection in a biological
CC   sample. The method of the invention provides a sensitive, accurate
CC   diagnostic and prognostic tool to provide adequate patient care and to
CC   prevent transmission of HCV by blood and by blood products, or by
CC   personal contact. Use of NS3/4a conformational epitope in combination
CC   with MEFA, provides a sensitive and reliable method for detecting early
CC   HCV seroconversion. Use of MEFA has the added advantages of decreasing
CC   masking problems, improving sensitivity in detecting antibodies by
CC   allowing a greater number of epitopes on a unit surface area of
CC   substrate, and improving substrate. Detection accuracy is increased and
CC   the incidence of false results is reduced because of the identification
CC   and the use of highly immunogenic HCV antigens which are present during
CC   the early stages of HCV seroconversion. The present amino acid sequence
CC   represents the multiple epitope fusion antigen (MEFA) 7.1 of the
CC   invention
XX
SQ   Sequence 1099 AA;
```

```
  Query Match             90.5%;  Score 4032;  DB 5;  Length 1099;
  Best Local Similarity   69.8%;  Pred. No. 5.4e-280;
  Matches  791;  Conservative    1;  Mismatches    3;  Indels  338;  Gaps    7;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSTR---------------------------------------------------- 71

Db         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy         72 ------------------------------------GCNCSIYPGHITGHRMAWKLGS 93

Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180

Qy         94 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLITS-------- 145

Db        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240

Qy        146 --------TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 197

Db        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300

Qy        198 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 257

Db        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360

Qy        258 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 317

Db        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420

Qy        318 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 377

Db        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480

Qy        378 GYTGDFDSVIDCNTC---------------------------------------------- 392
```

```
Db        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540

Qy        393 ------------------------------------------------------------ 392

Db        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600

Qy        393 ------------------------------------------------------------ 392

Db        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660

Qy        393 -----------------------------------ACSGKPAIIPDREVLYREFDEMEE 416
                                                  ||||||||||||||||||||||
Db        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSACSGKPAIIPDREVLYREFDEMEE 720

Qy        417 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 780

Qy        477 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 840

Qy        537 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 900

Qy        597 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 960

Qy        657 KRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVLATRKTSPIPKARRPEGRTWAQPGY 716
              ||||||||||||||||||||      ||||          ||||||||||||||||||
Db        961 KRNTNRRPQDVKFPGGGQIVG------RRGP----------PIPKARRPEGRTWAQPGY 1003

Qy        717 PWPLYGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPR 776
              ||||||||||||||||||||||||||||||||||||||||||||||||      ||||
Db       1004 PWPLYGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVG------RRGP- 1056

Qy        777 LGVLATRKTSPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 829
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1057 ----------PIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 1099


RESULT 4
ABG72262
ID   ABG72262 standard; protein; 1099 AA.
XX
AC   ABG72262;
XX
DT   06-MAR-2003  (first entry)
XX
DE   HCV multiple epitope fusion antigen 7.1 (MEFA 7.1).
XX
KW   Immunoassay solid support; Hepatitis C Virus type-1; HCV-1; HCV-2;
KW   NS3/4a conformational epitope; multiple epitope fusion antigen 7.1;
KW   MEFA 7.1; anti-HCV antibody; NS3/4a conformational antigen; HCV-3;
KW   HCV infection; Hepatitis C Virus type-2; Hepatitis C Virus type-3;
KW   mutant; mutein.
XX
OS   Hepatitis C virus type 1.
OS   Hepatitis C virus type 2.
OS   Hepatitis C virus type 3.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .156
FT                   /note= "Correspond to amino acids 1-156 of HCV-1 hSOD
FT                   superoxide dismutase)"
FT   Region          159. .176
FT                   /note= "Correspond to amino acids 303-320 of HCV-1 E1"
FT   Region          179. .199
FT                   /note= "Correspond to consensus sequence of amino acids
FT                   390-410 of HCV-1 E2 HVR"
FT   Region          200. .230
FT                   /note= "Correspond to consensus sequence of amino acids
FT                   384-414 of HCV-1 and HCV-2 E2 HVR"
FT   Region          231. .696
FT                   /note= "Correspond to amino acids 1193-1658 of HCV-1
FT                   helicase"
FT   Region          699. .745
FT                   /note= "Correspond to amino acids 1689-1735 of HCV-1 5-1-
FT                   1 epitope"
FT   Region          748. .794
FT                   /note= "Correspond to amino acids 1689-1735 of HCV-3 5-1-
FT                   1 epitope"
FT   Region          797. .843
FT                   /note= "Correspond to amino acids 1689-1735 of HCV-2 5-1-
FT                   1 epitope"
FT   Region          846. .881
FT                   /note= "Correspond to amino acids 1901-1936 of HCV-1
FT                   polypeptide C100"
FT   Region          884. .919
FT                   /note= "Correspond to amino acids 2278-2313 of HCV-1 NS5
FT                   region"
FT   Region          922. .957
FT                   /note= "Correspond to amino acids 2278-2313 of HCV-1 NS5
FT                   region"
FT   Region          958. .1028
FT                   /note= "Correspond to core region antigenic determinants
FT                   from amino acids 9-32, 39-42 and 64-88 of HCV-1 and amino
FT                   acids 67-84 of HCV-2"
FT   Region          1029. .1099
FT                   /note= "Correspond to core region antigenic determinants
FT                   from amino acids 9-32, 39-42 and 64-88 of HCV-1 and amino
FT                   acids 67-84 of HCV-2"
XX
PN   US2002146685-A1.
XX
PD   10-OCT-2002.
XX
PF   14-JUN-2001; 2001US-00881654.
XX
PR   15-JUN-2000; 2000US-0212082P.
PR   02-APR-2001; 2001US-0280811P.
PR   02-APR-2001; 2001US-0280867P.
XX
PA   (CHIE/) CHIEN D Y.
PA   (ARCA/) ARCANGEL P.
PA   (TAND/) TANDESKE L.
PA   (GEOR/) GEORGE-NASCIMENTO C.
PA   (COIT/) COIT D.
PA   (MEDI/) MEDINA-SELBY A.
XX
PI   Chien DY,  Arcangel P,  Tandeske L,  George-Nascimento C,  Coit D;
PI   Medina-Selby A;
XX
DR   WPI; 2003-147573/14.
DR   N-PSDB; ABX14411.
XX
PT   Immunoassay solid support for detecting Hepatitis C Virus infection in
PT   biological samples, comprises Hepatitis C Virus conformational epitope
PT   and multiple epitope fusion antigen.
XX
PS   Claim 25; Fig 5A-5F; 45pp; English.
XX
CC   The present invention relates to immunoassays comprising Hepatitis C
CC   Virus (HCV) NS3/4a conformational epitope and multiple epitope fusion
CC   antigen (MEFA), bound to a solid support. The NS3/4a epitope and/or the
CC   multiple epitope fusion antigen react with anti-HCV antibodies present in
CC   a biological sample from an HCV-infected individual. The immunoassays and
CC   methods of the invention are useful for detecting HCV infection in a
CC   biological sample. The inventive immunoassay solid support provides a
CC   sensitive and reliable method for detecting early HCV seroconversion. The
CC   assays can detect HCV infection caused by any six known genotypes of HCV.
CC   The use of the multiple epitope fusion proteins decreases masking
```

```
CC   problems, improves sensitivity in detecting antibodies by allowing a
CC   greater number of epitopes on a unit area of substrate, and improves
CC   selectivity. The present sequence represents HCV multiple epitope fusion
CC   antigen 7.1 (MEFA 7.1), a mutant HCV polyprotein derived from various
CC   regions of HCV type 1, 2, or 3 (HCV-1, HCV-2, or HCV-3) polypeptide
CC   sequences
XX
SQ   Sequence 1099 AA;

  Query Match             90.5%;  Score 4032;  DB 6;  Length 1099;
  Best Local Similarity   69.8%;  Pred. No. 5.4e-280;
  Matches  791;  Conservative    1;  Mismatches    3;  Indels  338;  Gaps    7;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSTR---------------------------------------------------- 71
              |||||||| :
Db         61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVV 120

Qy         72 ---------------------------------------GCNCSIYPGHITGHRMAWKLGS 93
                                                     |||||||||||||||||||||
Db        121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQNLNSGCNCSIYPGHITGHRMAWKLGS 180

Qy         94 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLITS-------- 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AARTTSGFVSLFAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQLIVDFIPVENLE 240

Qy        146 --------TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 197
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL 300

Qy        198 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTD 360

Qy        258 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPL 420

Qy        318 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMT 480

Qy        378 GYTGDFDSVIDCNTC---------------------------------------------- 392
              |||||||||||||||
Db        481 GYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFV 540

Qy        393 ------------------------------------------------------------ 392

Db        541 APGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVF 600

Qy        393 ------------------------------------------------------------ 392

Db        601 TGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPT 660

Qy        393 -----------------------------------ACSGKPAIIPDREVLYREFDEMEE 416
                                                  ||||||||||||||||||||||
Db        661 PLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSACSGKPAIIPDREVLYREFDEMEE 720

Qy        417 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CSQHLPYIEQGMMLAEQFKQKALGLSRGGKPAIVPDKEVLYQQYDEMEECSQAAPYIEQA 780

Qy        477 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 QVIAHQFKEKVLGLIDNDQVVVTPDKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKI 840

Qy        537 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 QGLLGILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYN 900

Qy        597 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 PPLVETWKKPDYEPPVVHGRSSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKT 960

Qy        657 KRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVLATRKTSPIPKARRPEGRTWAQPGY 716
              ||||||||||||||||||||      ||||           ||||||||||||||||||
Db        961 KRNTNRRPQDVKFPGGGQIVG------RRGP-----------PIPKARRPEGRTWAQPGY 1003

Qy        717 PWPLYGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPR 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||      ||||
Db       1004 PWPLYGNKDRRSTGKSWGKPGYPWPRKTKRNTNRRPQDVKFPGGGQIVG------RRGP- 1056

Qy        777 LGVLATRKTSPIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 829
                          ||||||||||||||||||||||||||||||||||||||||
Db       1057 ----------PIPKARRPEGRTWAQPGYPWPLYGNKDRRSTGKSWGKPGYPWP 1099


RESULT 5
AAW34481
ID   AAW34481 standard; protein; 1021 AA.
XX
AC   AAW34481;
XX
DT   25-MAR-2003  (revised)
DT   16-MAR-1998  (first entry)
XX
DE   HCV antigen combination pSOD/c200/core.
XX
KW   PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
KW   C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
KW   NS4.
XX
OS   Hepatitis C virus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .902
FT                   /note= "linker"
FT   Misc-difference 1. .154
FT                   /note= "hSOD fragment"
FT   Misc-difference 155. .159
FT                   /note= "linker"
FT   Misc-difference 160. .899
FT                   /note= "c200 (amino acids 1192-1931 of HCV polyprotein)"
FT   Misc-difference 903. .1021
FT                   /note= "c22 (amino acids 2-120 of HCV polyprotein)"
XX
PN   US5683864-A.
XX
PD   04-NOV-1997.
XX
PF   07-JUL-1992;   92US-00910760.
XX
PR   18-NOV-1987;   87US-00122714.
PR   30-DEC-1987;   87US-00139886.
PR   26-FEB-1988;   88US-00161072.
PR   06-MAY-1988;   88US-00191263.
PR   26-OCT-1988;   88US-00263584.
PR   14-NOV-1988;   88US-00271450.
PR   17-MAR-1989;   89US-00325338.
PR   20-APR-1989;   89US-00341334.
PR   21-APR-1989;   89US-00353896.
PR   18-MAY-1989;   89US-00355002.
PR   04-APR-1990;   90US-00504352.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kuo G,  Houghton M,  Choo Q;
XX
DR   WPI; 1997-548976/50.
DR   N-PSDB; AAT99982.
XX
PT   Combination of three hepatitis C virus antigens - used for detection of
```

```
PT   specific antibodies to diagnose infection.
XX
PS   Example 6; Col 59-68; 57pp; English.
XX
CC   This sequence represents a Hepatitis c virus (HCV) antigen combination of
CC   the invention. The HCV antigen combination comprises an antigen (Ag1)
CC   comprising the C domain (i.e. amino acids (aa) 1-120 of the HCV
CC   polyprotein), or its immunologically reactive fragment containing at
CC   least 8 aa. It also comprises two additional antigens from two different
CC   polyprotein domains, including at least 8 aa from the NS3, NS4, S or NS5
CC   domains of the polyprotein, corresponding, respectively, to aa 1050-1640;
CC   1640-2000; 120-400 and 2000-3011 of the HCV polyprotein. Alternatively,
CC   Ag1 contains at least 8 aa from the 1-122 or 9-177 aa regions of the HCV
CC   polyprotein. These antigen combinations are used diagnostically to detect
CC   anti-HCV antibodies, using any standard immunoassay format. These antigen
CC   combinations have a broader range of reactivity with antibodies than any
CC   antigen individually. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ   Sequence 1021 AA;

  Query Match             49.9%;  Score 2222;  DB 2;  Length 1021;
  Best Local Similarity   46.2%;  Pred. No. 3e-150;
  Matches  512;  Conservative   35;  Mismatches  111;  Indels  450;  Gaps   22;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSTRGCNCSIYPGHITGHRMAWKLGSAARTTSG---------FVSL------- 104
              |||||||| :      : |     |    ||:      |          :||
Db         61 AGPHFNPLSRK------HGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCII 114

Qy        105 ------------FAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQ----------- 141
                          |  :  |  || | :|   |:      | :||::
Db        115 GRTLVVHEKADDLGKGGNEESTK-TGNAGSRLACGVI-----GIAQNLEFGAVDFIPVEN 168

Qy        142 LITS------TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAA 195
              | |:      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 LETTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAA 228

Qy        196 TLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHS 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 TLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHS 288

Qy        256 TDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAI 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAI 348

Qy        316 PLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDAL 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 PLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDAL 408

Qy        376 MTGYTGDFDSVIDCNTC---------------------------------------- 392
              |||||||||||||||||
Db        409 MTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYR 468

Qy        393 ------------------------------------------------------------ 392

Db        469 FVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEG 528

Qy        393 ------------------------------------------------------------ 392

Db        529 VFTGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHG 588

Qy        393 ------------------------------------------------------------ 392

Db        589 PTPLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCV 648

Qy        393 ------ACSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----S 442
                     ||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db        649 VIVGRVVLSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTAS 708

Qy        443 RGGKPAIVPDKEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVV 497
```

```
                |  :  | |  :  :|: :      |       |:     :       :  |:     |
Db        709 RQAE-VIAPAVQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAV 765

Qy        498 VTP---DKEILYE----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQ 537
               :|    : :|:                ||      :    :  |: : ::|
Db        766 TSPLTTSQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGA 825

Qy        538 GLLG-----------------------------------------ILRRHVGPGEGAVQWMN 558
               |: |                                        ||||||||||||||||
Db        826 GVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMN 885

Qy        559 RLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRSS 618
               |||||||||||||                                          | ||
Db        886 RLIAFASRGNHVSP-----------------------------------------GNSS 903

Qy        619 RRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNTNRRPQDVKFPGGGQIVGG 678
                           ||      ||       :| ||||||||||||||||||||||||
Db        904 T-------------NP-------KPQ--------KKNKRNTNRRPQDVKFPGGGQIVGG 934

Qy        679 VYLLPRRGPRLGVLATRKTS----------PIPKARRPEGRTWAQPGYPWPLYGNK---- 724
               |||||||||||| ||||||          |||||||||||||||||||||:
Db        935 VYLLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGW 994

Qy        725 -----DRRSTGKSWGKPGYPWPRKTKRN 747
                    | :  |||      ||:  ||
Db        995 AGWLLSPRGSRPSWGPTD---PRRRSRN 1019
```

```
RESULT 6
AAW40039
ID   AAW40039 standard; protein; 1021 AA.
XX
AC   AAW40039;
XX
DT   26-MAY-1998  (first entry)
XX
DE   Fusion protein c200/c22.
XX
KW   Hepatitis C virus C domain; HCV; immunological activity; c200/c22;
KW   NS3 domain; NS4 domain; S domain; NS5 domain; fusion protein.
XX
OS   Synthetic.
OS   Hepatitis virus.
XX
PN   US5712087-A.
XX
PD   27-JAN-1998.
XX
PF   12-MAY-1995;   95US-00440519.
XX
PR   04-APR-1990;   90US-00504352.
PR   07-JUL-1992;   92US-00910760.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kuo G,  Houghton M,  Choo Q;
XX
DR   WPI; 1998-119973/11.
DR   N-PSDB; AAV09990.
XX
PT   Immunoassays for hepatitis C virus antibodies - using combinations of
PT   antigenic fragments of HCV polyprotein.
XX
PS   Example 6; Fig 4; 59pp; English.
XX
CC   This sequence represents a fusion protein constructed from the hepatitis
CC   C virus core domain (which is situated at the carboxy terminus of the
CC   fusion protein) and a c200 construct (a fusion of the NS3 and NS3
CC   domains). This protein used in the construction of novel combinations of
CC   HCV antigens that have a broader range of immunological activity than any
CC   single HCV antigen. An example of such an antigen given in this
CC   specification comprises a first antigen containing at least 8 amino acids
CC   of the C domain of the HCV polyprotein and a second antigen comprising at
```

```
CC   least 8 amino acids of the NS3 domain, the NS4 domain, the S domain or
CC   the NS5 domain of the HCV polyprotein in the form of a fusion protein, a
CC   physical mixture or bound to a solid matrix
XX
SQ   Sequence 1021 AA;

  Query Match             49.9%;  Score 2222;  DB 2;  Length 1021;
  Best Local Similarity   46.2%;  Pred. No. 3e-150;
  Matches  512;  Conservative   35;  Mismatches  111;  Indels  450;  Gaps   22;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSTRGCNCSIYPGHITGHRMAWKLGSAARTTSG---------FVSL------- 104
              |||||||| :      : |     |    ||:      |          :||
Db         61 AGPHFNPLSRK------HGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCII 114

Qy        105 ------------FAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQ----------- 141
                          |  :  |  || | :|   |:      | :||::
Db        115 GRTLVVHEKADDLGKGGNEESTK-TGNAGSRLACGVI-----GIAQNLEFGAVDFIPVEN 168

Qy        142 LITS------TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAA 195
              | |:      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 LETTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAA 228

Qy        196 TLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHS 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 TLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHS 288

Qy        256 TDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAI 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAI 348

Qy        316 PLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDAL 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 PLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDAL 408

Qy        376 MTGYTGDFDSVIDCNTC---------------------------------------------- 392
              |||||||||||||||||
Db        409 MTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYR 468

Qy        393 ------------------------------------------------------------ 392

Db        469 FVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEG 528

Qy        393 ------------------------------------------------------------ 392

Db        529 VFTGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHG 588

Qy        393 ------------------------------------------------------------ 392

Db        589 PTPLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCV 648

Qy        393 ------ACSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----S 442
                     ||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db        649 VIVGRVVLSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTAS 708

Qy        443 RGGKPAIVPDKEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVV 497
              |  :  | |   :  :|: :       |         |:     :        :  |:     |
Db        709 RQAE-VIAPAVQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAV 765

Qy        498 VTP---DKEILYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQ 537
               :|    : :|:                  ||       :    :  |: : ::|
Db        766 TSPLTTSQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGA 825

Qy        538 GLLG----------------------------------------ILRRHVGPGEGAVQWMN 558
              |: |                                        ||||||||||||||||
Db        826 GVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMN 885

Qy        559 RLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRSS 618
              ||||||||||||                                            | ||
Db        886 RLIAFASRGNHVSP----------------------------------------GNSS 903

Qy        619 RRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNTNRRPQDVKFPGGGQIVGG 678
                              ||       ||         :| ||||||||||||||||||||
Db        904 T--------------NP-------KPQ--------KKNKRNTNRRPQDVKFPGGGQIVGG 934

Qy        679 VYLLPRRGPRLGVLATRKTS----------PIPKARRPEGRTWAQPGYPWPLYGNK---- 724
              |||||||||||||| ||||||          |||||||||||||||||||||||:
Db        935 VYLLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGW 994

Qy        725 -----DRRSTGKSWGKPGYPWPRKTKRN 747
                  | :  |||      ||:  ||
Db        995 AGWLLSPRGSRPSWGPTD---PRRRSRN 1019


RESULT 7
AAE22050
ID   AAE22050 standard; protein; 1021 AA.
XX
AC   AAE22050;
XX
DT   16-JUL-2002  (first entry)
XX
DE   pSOD/c200/core expression plasmid protein.
XX
KW   Hepatitis C virus; HCV; antigen; C domain; polyprotein; NS3 domain;
KW   NS4 domain; S domain; NS5 domain; pSOD/c200/core plasmid.
XX
OS   Hepatitis C virus.
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .154
FT                   /note= "hSOD"
FT   Region          155. .159
FT                   /note= "Linker region"
FT   Region          160. .899
FT                   /note= "HCV c200"
FT   Region          900. .902
FT                   /note= "Linker region"
FT   Region          903. .1021
FT                   /note= "HCV c22"
XX
PN   US6312889-B1.
XX
PD   06-NOV-2001.
XX
PF   12-MAY-1995;   95US-00440549.
XX
PR   04-APR-1990;   90US-00504352.
PR   07-JUL-1992;   92US-00910760.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Houghton M,  Choo Q,  Kuo G;
XX
DR   WPI; 2002-040268/05.
DR   N-PSDB; AAD35044.
XX
PT   Combination of hepatitis C viral (HCV) antigens, useful in improved
PT   immunoassay for detecting HCV antibodies.
XX
PS   Example 6; Fig 4; 58pp; English.
XX
CC   The invention relates to combination of hepatitis C viral (HCV) antigens
CC   that have a broader range of immunological reactivity than any single HCV
CC   antigen. The combinations consist of an antigen from the C domain of the
CC   HCV polyprotein, and at least one additional HCV antigen from either the
CC   NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
CC   the form of fusion protein, a simple physical mixture, or the individual
CC   antigens commonly bound to a solid matrix. The combinations of antigens
CC   provides broad range immunoassays for anti-HCV antibodies. The invention
CC   therefore provides a method for detecting antibodies to HCV in a mammal
```

```
CC   suspected of containing such antibodies. The present sequence is a
CC   protein encoded by pSOD/c200/core expression plasmid DNA containing HCV
CC   coding sequence
XX
SQ   Sequence 1021 AA;

  Query Match             49.9%;  Score 2222;  DB 5;  Length 1021;
  Best Local Similarity   46.2%;  Pred. No. 3e-150;
  Matches  512;  Conservative   35;  Mismatches  111;  Indels  450;  Gaps   22;

Qy          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTS 60

Qy         61 AGPHFNPLSTRGCNCSIYPGHITGHRMAWKLGSAARTTSG---------FVSL------- 104
              |||||||||  :      : |     |    ||:       |           :||
Db         61 AGPHFNPLSRK------HGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCII 114

Qy        105 ------------FAPGAKQNETHVTGGAAARTTSGLTSLFSPGASQNIQ----------- 141
                          |  :  |  || | :|   |:      | :||::
Db        115 GRTLVVHEKADDLGKGGNEESTK-TGNAGSRLACGVI-----GIAQNLEFGAVDFIPVEN 168

Qy        142 LITS------TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAA 195
              | |:      ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 LETTMRSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAA 228

Qy        196 TLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHS 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 TLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHS 288

Qy        256 TDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAI 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAI 348

Qy        316 PLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDAL 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 PLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDAL 408

Qy        376 MTGYTGDFDSVIDCNTC---------------------------------------------- 392
              |||||||||||||||||
Db        409 MTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYR 468

Qy        393 ------------------------------------------------------------ 392

Db        469 FVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEG 528

Qy        393 ------------------------------------------------------------ 392

Db        529 VFTGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHG 588

Qy        393 ------------------------------------------------------------ 392

Db        589 PTPLLYRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCV 648

Qy        393 ------ACSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----S 442
                     ||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db        649 VIVGRVVLSGKPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTAS 708

Qy        443 RGGKPAIVPDKEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVV 497
              |  :  | |  :  :|: :      |         |:     :       :  |:     |
Db        709 RQAE-VIAPAVQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAV 765

Qy        498 VTP---DKEILYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQ 537
              :|    : :|:                  ||      :    |: : ::|
Db        766 TSPLTTSQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGA 825

Qy        538 GLLG-----------------------------------------ILRRHVGPGEGAVQWMN 558
              |: |                                         ||||||||||||||||
Db        826 GVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMN 885

Qy        559 RLIAFASRGNHVSPTHYVPSRSRRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRSS 618
              ||||||||||||||                                        | ||
Db        886 RLIAFASRGNHVSP-----------------------------------------GNSS 903

Qy        619 RRFAQALPVWARPDYNPPLVETWKKPDYEPPVVHGRKTKRNTNRRPQDVKFPGGGQIVGG 678
                                  ||         ||          :| ||||||||||||||||||||
Db        904 T--------------NP-------KPQ--------KKNKRNTNRRPQDVKFPGGGQIVGG 934

Qy        679 VYLLPRRGPRLGVLATRKTS----------PIPKARRPEGRTWAQPGYPWPLYGNK---- 724
              |||||||||||||| ||||||          |||||||||||||||||||||||:
Db        935 VYLLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGW 994

Qy        725 -----DRRSTGKSWGKPGYPWPRKTKRN 747
                    | :  |||      ||:  ||
Db        995 AGWLLSPRGSRPSWGPTD---PRRRSRN 1019


RESULT 8
AAP92041
ID   AAP92041 standard; protein; 1766 AA.
XX
AC   AAP92041;
XX
DT   25-MAR-2003  (revised)
DT   02-MAR-1990  (first entry)
XX
DE   Hepatitis C virus (HCV) cDNA inserts in clones 14i, 11b, 7f, 7e, 8h, 33c,
DE   40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c.
XX
KW   Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH).
XX
OS   Hepatitis C virus.
XX
PN   EP318216-A.
XX
PD   31-MAY-1989.
XX
PF   18-NOV-1988;   88EP-00310922.
XX
PR   18-NOV-1987;   87US-00122714.
PR   30-DEC-1987;   87US-00139886.
PR   26-FEB-1988;   88US-00161072.
PR   06-MAY-1988;   88US-00191263.
PR   26-OCT-1988;   88US-00263584.
PR   14-NOV-1988;   88US-00271450.
XX
PA   (CHIR ) CHIRON CORP.
PA   (CHIR ) CHIRON CORP.
XX
PI   Houghton M,  Choo QL,  Kuo G;
XX
DR   WPI; 1989-159274/22.
DR   N-PSDB; AAN92097.
XX
PT   Purified hepatitis C virus - and associated nucleic acids and
PT   polypeptide(s).
XX
PS   Claim 13; Fig 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
XX
CC   It is the sequence encoded in the open reading frame of hepatitis C virus
CC   cDNA inserts in clones 14i,m 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81,
CC   32, 33b, 25c, 14c, 8f, a33f, 33g and 39c. It is antigenic and could be
CC   used in immunoassay reagents and vaccines and to generate antibodies
CC   useful in diagnosis and passive immunotherapy for HCV infection/non-A,
CC   non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.) (Updated
CC   on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 1766 AA;

  Query Match             36.5%;  Score 1627.5;  DB 1;  Length 1766;
  Best Local Similarity   36.6%;  Pred. No. 2.7e-107;
  Matches  398;  Conservative   26;  Mismatches   81;  Indels  581;  Gaps   15;

Qy        146 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 554
```

```
Qy        206 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 265
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        555 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYHIIICDECHSTDATSILGIG 614

Qy        266 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 674

Qy        326 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 734

Qy        386 VIDCNTC---------------------------------------------------- 392
              |||||||
Db        735 VIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPGERPSG 794

Qy        393 ---------------------------------------------------------- 392

Db        795 MFDSSVLCECYDEGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGLTHIDA 854

Qy        393 ---------------------------------------------------------- 392

Db        855 HFLSQTKQSGENLPYLVAYQATVWARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLYRLGA 914

Qy        393 ----------------------------------------------------------ACSG 396
                                                                           ||
Db        915 VQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRVVLSG 974

Qy        397 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRGGKPAIVPD 452
              |||||||||||||||||||||||||||||||||||||||||||||    ||  :  | |
Db        975 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQAE-VIAPA 1033

Qy        453 KEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVTP---DKEI 504
               :  :|: :       |        |:       :       :  |:      | :|    : :
Db       1034 VQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTSPLTTSQTL 1091

Qy        505 LYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGLLG------ 541
              |:                 ||       :    :  |: : ::|      |: |
Db       1092 LFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGVAGALVAFK 1151

Qy        542 ---------------------------------ILRRHVGPGEGAVQWMNRLIAFASRGN 568
                                                |||||||||||||||||||||||||
Db       1152 IMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGN 1211

Qy        569 HVSPTHYVPS---------------------------------------------------- 578
              |||||||||
Db       1212 HVSPTHYVPESDAAARVTAILSSLTVTQLLRRLHQWISSECTTPCSGSWLRDIWDWICEV 1271

Qy        579 ---------------------------------------------------------- 578

Db       1272 LSDFKTWLKAKLMPQLPGIPFVSCQRGYKGVWRVDGIMHTRCHCGAEITGHVKNGTMRIV 1331

Qy        579 ---------------------------------------------------------- 578

Db       1332 GPRTCRNMWSGTFPINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVPSPEFFTELDG 1391

Qy        579 -RSRRFA----------------------QALPVWARPDY------------------- 595
               |  |||                      ||    ||
Db       1392 VRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPDVAVLTSMLTDPSHITAEAAG 1451

Qy        596 ------NPP------------------------------LVET---WKK------------ 605
                   :||                             |:|    |::
Db       1452 RRLARGSPPSVASSSASQLSAPSLKATCTANHDSPDAELIEANLLWRQEMGGNITRVESE 1511

Qy        606 ------PDYEPPVVHG------------RSSRRFAQALPVWARPDYNPPLVETWKKPDYE 647
                     ::| |               | ||||||||||||||||||||||||||||||
Db       1512 NKVVILDSFDPLVAEEDEREISVPAEILRKSRRFAQALPVWARPDYNPPLVETWKKPDYE 1571

Qy        648 PPVVHG 653
              ||||||
Db       1572 PPVVHG 1577
```

RESULT 9
AAP90164
ID AAP90164 standard; protein; 2261 AA.
XX
AC AAP90164;
XX
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Peptide encoded by composite hepatitis C virus cDNA.
XX
KW Hepatitis C virus; clone 12f; clone 15e; probe; vaccine.
XX
OS Pan troglodytes.
XX
PN GB2212511-A.
XX
PD 26-JUL-1989.
XX
PF 18-NOV-1988; 88GB-00027024.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 26-OCT-1988; 88US-00263584.
XX
PA (CHIR) CHIRON CORP.
XX
PI Houghton M, Choo QL, Kuo G;
XX
DR WPI; 1989-215054/30.
DR N-PSDB; AAN90331.
XX
PT Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
PS Disclosure; Fig 32; 30pp; English.
XX
CC The sequence is the peptide encoded by the composite hepatitis C virus
CC (HCV) cDNA of AAN90331. The polypeptides are used to diagnose HCV-induced
CC NANBH, to raise antibodies for immunoassay or treatment, or to produce
CC vaccines. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 2261 AA;

Query Match 36.5%; Score 1624.5; DB 1; Length 2261;
Best Local Similarity 36.1%; Pred. No. 6.2e-107;
Matches 399; Conservative 26; Mismatches 80; Indels 601; Gaps 15;

```
Qy        146 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 645

Qy        206 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 705

Qy        266 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 765

Qy        326 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 825

Qy        386 VIDCNTC---------------------------------------------------- 392
              |||||||
Db        826 VIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPGERPSG 885

Qy        393 ---------------------------------------------------------- 392
```

```
Db        886 MFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGLTHIDA 945

Qy        393 ------------------------------------------------------------ 392

Db        946 HFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLYRLGA 1005

Qy        393 ------------------------------------------------------------ACSG 396
                                                                             ||
Db       1006 VQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRVVLSG 1065

Qy        397 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRGGKPAIVPD 452
              ||||||||||||||||||||||||||||||||||||||||||||    || :  | |
Db       1066 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQAE-VIAPA 1124

Qy        453 KEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVTP---DKEI 504
               :  :|: :      |       |:     :       :  |:     | :|    : :
Db       1125 VQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTSPLTTSQTL 1182

Qy        505 LYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGLLG------ 541
               |:                 ||       :    :  |: : ::|      |: |
Db       1183 LFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGVAGALVAFK 1242

Qy        542 --------------------------------ILRRHVGPGEGAVQWMNRLIAFASRGN 568
                                               |||||||||||||||||||||||||||
Db       1243 IMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGN 1302

Qy        569 HVSPTHYVPS---------------------------------------------------- 578
               |||||||||
Db       1303 HVSPTHYVPESDAAARVTAILSSLTVTQLLRRLHQWISSECTTPCSGSWLRDIWDWICEV 1362

Qy        579 ------------------------------------------------------------ 578

Db       1363 LSDFKTWLKAKLMPQLPGIPFVSCQRGYKGVWRVDGIMHTRCHCGAEITGHVKNGTMRIV 1422

Qy        579 ------------------------------------------------------------ 578

Db       1423 GPRTCRNMWSGTFPINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTT 1482

Qy        579 ---------------------RSRRFA----------------------QALPVWARPD 594
                                   |  |||                      ||    ||
Db       1483 DNLKCPCQVPSPEFFTELDGVRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPD 1542

Qy        595 Y-------------------------NPP-----------------------------LV 600
                                       :||                             |:
Db       1543 VAVLTSMLTDPSHITAEAAGRRLARGSPPSVASSSASQLSAPSLKATCTANHDSPDAELI 1602

Qy        601 ET---WKK-------------------PDYEPPVVHG------------RSSRRFAQALPV 627
               |    |::                    ::| |              | ||||||||||
Db       1603 EANLLWRQEMGGNITRVESENKVVILDSFDPLVAEEDEREISVPAEILRKSRRFAQALPV 1662

Qy        628 WARPDYNPPLVETWKKPDYEPPVVHG 653
              |||||||||||||||||||||||||
Db       1663 WARPDYNPPLVETWKKPDYEPPVVHG 1688
```

RESULT 10
AAP92050
ID AAP92050 standard; protein; 2436 AA.
XX
AC AAP92050;
XX
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX
DE Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones K9
DE -1 through 15e.
XX
KW Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH).
XX
OS Hepatitis C virus.
XX
PN EP318216-A.
XX
PD 31-MAY-1989.
XX
PF 18-NOV-1988; 88EP-00310922.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 06-MAY-1988; 88US-00191263.
PR 26-OCT-1988; 88US-00263584.
PR 14-NOV-1988; 88US-00271450.
XX
PA (CHIR) CHIRON CORP.
PA (CHIR) CHIRON CORP.
XX
PI Houghton M, Choo QL, Kuo G;
XX
DR WPI; 1989-159274/22.
DR N-PSDB; AAN92106.
XX
PT Purified hepatitis C virus - and associated nucleic acids and
PT polypeptide(s).
XX
PS Claim 13; Fig 47-1-47-8; 139pp; English.
XX
CC It is the sequence encoded in the open reading frame of hepatitis C virus
CC (HCV) cDNA inserts in clones K9-1 through 15e. It is antigenic and could
CC be used in immunoassay reagents and vaccines and to generate antibodies
CC useful in diagnosis and passive immunotherapy for HCV infection/non-A,
CC non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.) (Updated
CC on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 2436 AA;

Query Match 36.5%; Score 1624.5; DB 1; Length 2436;
Best Local Similarity 36.1%; Pred. No. 6.9e-107;
Matches 399; Conservative 26; Mismatches 80; Indels 601; Gaps 15;

```
Qy        146 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 205
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 820

Qy        206 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 265
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 880

Qy        266 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 325
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        881 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 940

Qy        326 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 385
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 1000

Qy        386 VIDCNTC---------------------------------------------------- 392
              |||||||
Db       1001 VIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPGERPSG 1060

Qy        393 ------------------------------------------------------------ 392

Db       1061 MFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGLTHIDA 1120

Qy        393 ------------------------------------------------------------ 392

Db       1121 HFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLYRLGA 1180

Qy        393 ------------------------------------------------------------ACSG 396
                                                                             ||
Db       1181 VQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRVVLSG 1240

Qy        397 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRGGKPAIVPD 452
              ||||||||||||||||||||||||||||||||||||||||||||    || :  | |
Db       1241 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQAE-VIAPA 1299
```

```
Qy        453 KEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVTP---DKEI 504
Db       1300 VQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTSPLTTSQTL 1357

Qy        505 LYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGLLG------ 541
Db       1358 LFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGVAGALVAFK 1417

Qy        542 --------------------------------ILRRHVGPGEGAVQWMNRLIAFASRGN 568
Db       1418 IMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGN 1477

Qy        569 HVSPTHYVPS---------------------------------------------------- 578
Db       1478 HVSPTHYVPESDAAARVTAILSSLTVTQLLRRLHQWISSECTTPCSGSWLRDIWDWICEV 1537

Qy        579 ------------------------------------------------------------ 578
Db       1538 LSDFKTWLKAKLMPQLPGIPFVSCQRGYKGVWRVDGIMHTRCHCGAEITGHVKNGTMRIV 1597

Qy        579 ------------------------------------------------------------ 578
Db       1598 GPRTCRNMWSGTFPINAYTTGPCTPLPAPNYTFALWRVSAEEYVEVRQVGDFHYVAGMIT 1657

Qy        579 ---------------------RSRRFA----------------------QALPVWARPD 594
Db       1658 DNLKCPCQVPSPEFFTELDGVRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPD 1717

Qy        595 Y-------------------------NPP---------------------------LV 600
Db       1718 VAVLTSMLTDPSHITAEAAGRRLARGSPPSVASSSASQLSAPSLKATCTANHDSPDAELI 1777

Qy        601 ET---WKK-------------------PDYEPPVVHG------------RSSRRFAQALPV 627
Db       1778 EANLLWRQEMGGNITRVESENKVVILDSFDPLVAEEDEREISVPAEILRKSRRFAQALPV 1837

Qy        628 WARPDYNPPLVETWKKPDYEPPVVHG 653
Db       1838 WARPDYNPPLVETWKKPDYEPPVVHG 1863
```

RESULT 11
AAP90288
ID AAP90288 standard; protein; 2436 AA.
XX
AC AAP90288;
XX
DT 25-MAR-2003 (revised)
DT 19-JUL-2001 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Peptide encoded by composite hepatitis C cDNA.
XX
KW Hepatitis C virus; clone 15e; clone k9-1; probe; vaccine.
XX
OS Pan troglodytes.
XX
PN GB2212511-A.
XX
PD 26-JUL-1989.
XX
PF 18-NOV-1988; 88GB-00027024.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 26-OCT-1988; 88US-00263584.
XX
PA (CHIR) CHIRON CORP.
XX
PI Houghton M, Choo QL, Kuo G;
XX
DR WPI; 1989-215054/30.
DR N-PSDB; AAN90336.
XX
PT Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
PS Disclosure; Fig 47-1 to 47-8; 30pp; English.
XX
CC The sequence is the peptide encoded by the composite hepatitis C virus
CC (HCV) cDNA of AAN90336. The polypeptides are used to diagnose HCV-induced
CC NANBH, to raise antibodies for immunoassay or treatment, or to produce
CC vaccines. (N.B. This record was resubmitted to correct errors in the
CC sequence.) (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 2436 AA;

Query Match 36.5%; Score 1624.5; DB 1; Length 2436;
Best Local Similarity 36.1%; Pred. No. 6.9e-107;
Matches 399; Conservative 26; Mismatches 80; Indels 601; Gaps 15;

```
Qy        146 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 205
Db        761 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 820

Qy        206 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 265
Db        821 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 880

Qy        266 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 325
Db        881 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 940

Qy        326 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 385
Db        941 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 1000

Qy        386 VIDCNTC---------------------------------------------------- 392
Db       1001 VIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPGERPSG 1060

Qy        393 ------------------------------------------------------------ 392
Db       1061 MFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGLTHIDA 1120

Qy        393 ------------------------------------------------------------ 392
Db       1121 HFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLYRLGA 1180

Qy        393 ---------------------------------------------------------ACSG 396
Db       1181 VQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRVVLSG 1240

Qy        397 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRGGKPAIVPD 452
Db       1241 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQAE-VIAPA 1299

Qy        453 KEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVTP---DKEI 504
Db       1300 VQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTSPLTTSQTL 1357

Qy        505 LYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGLLG------ 541
Db       1358 LFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGVAGALVAFK 1417

Qy        542 --------------------------------ILRRHVGPGEGAVQWMNRLIAFASRGN 568
Db       1418 IMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGN 1477

Qy        569 HVSPTHYVPS---------------------------------------------------- 578
Db       1478 HVSPTHYVPESDAAARVTAILSSLTVTQLLRRLHQWISSECTTPCSGSWLRDIWDWICEV 1537

Qy        579 ------------------------------------------------------------ 578
```

```
Db       1538 LSDFKTWLKAKLMPQLPGIPFVSCQRGYKGVWRVDGIMHTRCHCGAEITGHVKNGTMRIV 1597

Qy        579 ------------------------------------------------------------ 578

Db       1598 GPRTCRNMWSGTFPINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTT 1657

Qy        579 ---------------------RSRRFA----------------------QALPVWARPD 594
                                   |  |||                      ||    ||
Db       1658 DNLKCPCQVPSPEFFTELDGVRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPD 1717

Qy        595 Y-------------------------NPP----------------------------LV 600
                                       :||                            |:
Db       1718 VAVLTSMLTDPSHITAEAAGRRLARGSPPSVASSSASQLSAPSLKATCTANHDSPDAELI 1777

Qy        601 ET---WKK-------------------PDYEPPVVHG------------RSSRRFAQALPV 627
               |    |::                  ::| |               | ||||||||||
Db       1778 EANLLWRQEMGGNITRVESENKVVILDSFDPLVAEEDEREISVPAEILRKSRRFAQALPV 1837

Qy        628 WARPDYNPPLVETWKKPDYEPPVVHG 653
              |||||||||||||||||||||||||
Db       1838 WARPDYNPPLVETWKKPDYEPPVVHG 1863
```

RESULT 12
AAB18540
ID AAB18540 standard; protein; 2772 AA.
XX
AC AAB18540;
XX
DT 15-JAN-2001 (first entry)
XX
DE Protein encoded by a cDNA compiled Hepatitis C virus cDNA clones.
XX
KW Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;
KW viral infectivity; viral replication.
XX
OS Hepatitis C virus.
XX
PN EP1034785-A2.
XX
PD 13-SEP-2000.
XX
PF 16-MAR-1990; 2000EP-00109602.
XX
PR 17-MAR-1989; 89US-00325338.
PR 20-APR-1989; 89US-00341334.
PR 18-MAY-1989; 89US-00355002.
PR 16-MAR-1990; 90EP-00302866.
XX
PA (CHIR) CHIRON CORP.
XX
PI Houghton M, Choo Q, Kuo G;
XX
DR WPI; 2000-566891/53.
DR N-PSDB; AAA75296.
XX
PT Novel composition comprising a hepatitis C virus antisense polynucleotide
PT which is complementary to or corresponds to a sense strand of the virus
PT genome, and selectively hybridizes to it.
XX
PS Example; Fig 16; 75pp; English.
XX
CC The specification describes a pharmaceutical composition which comprises
CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
CC characterized by a positive stranded RNA genome which has 40% homology at
CC the polypeptide level to a HCV polyprotein. The antisense polynucleotide
CC binds to cellular polynucleotides which enhance and/or are required for
CC viral infectivity, replicative ability or chronicity. The antisense
CC polynucleotides may also be designed to bind with high specificity, to be
CC of increased stability, to be stable and to have low toxicity. The
CC composition also comprises an agent which causes viral RNA to be
CC inactive. The composition is used for preventing HCV replication in a
CC system. The present sequence is encoded by a novel HCV cDNA sequence,
CC which is used in the course of the invention
XX
SQ Sequence 2772 AA;

Query Match 36.5%; Score 1624.5; DB 3; Length 2772;
Best Local Similarity 36.1%; Pred. No. 8.2e-107;
Matches 399; Conservative 26; Mismatches 80; Indels 601; Gaps 15;

```
Qy        146 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 1156

Qy        206 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1157 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 1216

Qy        266 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1217 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 1276

Qy        326 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1277 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 1336

Qy        386 VIDCNTC---------------------------------------------------- 392
              |||||||
Db       1337 VIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPGERPSG 1396

Qy        393 ------------------------------------------------------------ 392

Db       1397 MFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGLTHIDA 1456

Qy        393 ------------------------------------------------------------ 392

Db       1457 HFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLYRLGA 1516

Qy        393 ------------------------------------------------------------ACSG 396
                                                                           ||
Db       1517 VQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRVVLSG 1576

Qy        397 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRGGKPAIVPD 452
              |||||||||||||||||||||||||||||||||||||||||||||    ||  :  | |
Db       1577 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQAE-VIAPA 1635

Qy        453 KEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVTP---DKEI 504
               :  :|: :     |         |:    :       :  |:     | :|    : :
Db       1636 VQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTSPLTTSQTL 1693

Qy        505 LYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGLLG------ 541
              |:                  ||       :   :  |: : ::|     |: |
Db       1694 LFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGVAGALVAFK 1753

Qy        542 ---------------------------------ILRRHVGPGEGAVQWMNRLIAFASRGN 568
                                                ||||||||||||||||||||||||||
Db       1754 IMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGN 1813

Qy        569 HVSPTHYVPS---------------------------------------------------- 578
              |||||||||
Db       1814 HVSPTHYVPESDAAARVTAILSSLTVTQLLRRLHQWISSECTTPCSGSWLRDIWDWICEV 1873

Qy        579 ------------------------------------------------------------ 578

Db       1874 LSDFKTWLKAKLMPQLPGIPFVSCQRGYKGVWRVDGIMHTRCHCGAEITGHVKNGTMRIV 1933

Qy        579 ------------------------------------------------------------ 578

Db       1934 GPRTCRNMWSGTFPINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTT 1993

Qy        579 ---------------------RSRRFA----------------------QALPVWARPD 594
                                   |  |||                      ||    ||
Db       1994 DNLKCPCQVPSPEFFTELDGVRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPD 2053
```

```
Qy        595 Y-------------------------NPP-----------------------------LV 600
                                        :||                             |:
Db       2054 VAVLTSMLTDPSHITAEAAGRRLARGSPPSVASSSASQLSAPSLKATCTANHDSPDAELI 2113

Qy        601 ET---WKK-------------------PDYEPPVVHG------------RSSRRFAQALPV 627
              |    |::                    ::| |               | ||||||||||
Db       2114 EANLLWRQEMGGNITRVESENKVVILDSFDPLVAEEDEREISVPAEILRKSRRFAQALPV 2173

Qy        628 WARPDYNPPLVETWKKPDYEPPVVHG 653
              |||||||||||||||||||||||||
Db       2174 WARPDYNPPLVETWKKPDYEPPVVHG 2199
```

```
RESULT 13
AAY14975
ID   AAY14975 standard; protein; 2955 AA.
XX
AC   AAY14975;
XX
DT   20-MAR-2003  (revised)
DT   08-NOV-1999  (first entry)
XX
DE   Amino acid sequence of HCV-1 ORF.
XX
KW   Hepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH;
KW   HCV infection; vaccine.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 441
FT                   /note= "encoded by TT"
FT   Misc-difference 461
FT                   /note= "encoded by CCCC"
XX
PN   EP939128-A2.
XX
PD   01-SEP-1999.
XX
PF   17-SEP-1990;   99EP-00101746.
XX
PR   15-SEP-1989;   89US-00408045.
PR   21-DEC-1989;   89US-00456142.
PR   17-SEP-1990;   90EP-00310149.
XX
PA   (OYAA/) OYA A.
PA   (CHIR ) CHIRON CORP.
XX
PI   Miyamura T,  Saito I,  Houghton M,  Weiner AJ,  Han J,  Kolberg JA;
PI   Cha T,  Irvine BD;
XX
DR   WPI; 1999-480843/41.
DR   N-PSDB; AAZ07656.
XX
PT   New Hepatitis C Virus isolates, useful for diagnosis of hepatitis
PT   infections and development of vaccines.
XX
PS   Disclosure; Fig 12; 132pp; English.
XX
CC   The invention provides two new isolates of hepatitis C virus (HCV), J1
CC   and J7. These two isolates comprise nucleotide and amino acid sequences
CC   that are distinct from the HCV isolate HCV-1. The nucleotide sequences
CC   may be used to detect non-A, non-B HCV (NANBH) polynucleotides by
CC   hybridisation for diagnosis of NANBH infections. They may also be used to
CC   screen blood donors, donated blood and blood products for this infection.
CC   The isolates may also be used to isolate other naturally occurring
CC   variants of the virus. The polypeptides may be used as a vaccine for
CC   administration to patients to protect against infection with NANBH. The
CC   present sequence represents the amino acid sequence of HCV-1 ORF.
CC   (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
CC   correct PR field.)
XX
SQ   Sequence 2955 AA;
```

```
  Query Match           36.5%;  Score 1624.5;  DB 2;  Length 2955;
  Best Local Similarity  36.1%;  Pred. No. 9e-107;
  Matches  399;  Conservative   26;  Mismatches   80;  Indels  601;  Gaps   15;

Qy        146 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 205
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1211 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 1270

Qy        206 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 265
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1271 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 1330

Qy        266 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 325
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1331 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 1390

Qy        326 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 385
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1391 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 1450

Qy        386 VIDCNTC---------------------------------------------------- 392
              |||||||
Db       1451 VIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPGERPSG 1510

Qy        393 ---------------------------------------------------------- 392

Db       1511 MFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGLTHIDA 1570

Qy        393 ---------------------------------------------------------- 392

Db       1571 HFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLYRLGA 1630

Qy        393 ------------------------------------------------------ACSG 396
                                                                     ||
Db       1631 VQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRVVLSG 1690

Qy        397 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRGGKPAIVPD 452
              |||||||||||||||||||||||||||||||||||||||||||||    ||  :  | |
Db       1691 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQAE-VIAPA 1749

Qy        453 KEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVTP---DKEI 504
              :  :|: :      |        |:     :       :  |:     | :|    : :
Db       1750 VQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTSPLTTSQTL 1807

Qy        505 LYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGLLG------ 541
              |:                 ||     :   :  |: : ::|     |: |
Db       1808 LFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGVAGALVAFK 1867

Qy        542 --------------------------------ILRRHVGPGEGAVQWMNRLIAFASRGN 568
                                              |||||||||||||||||||||||||||
Db       1868 IMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGN 1927

Qy        569 HVSPTHYVPS---------------------------------------------------- 578
              |||||||||
Db       1928 HVSPTHYVPESDAAARVTAILSSLTVTQLLRRLHQWISSECTTPCSGSWLRDIWDWICEV 1987

Qy        579 ---------------------------------------------------------- 578

Db       1988 LSDFKTWLKAKLMPQLPGIPFVSCQRGYKGVWRVDGIMHTRCHCGAEITGHVKNGTMRIV 2047

Qy        579 ---------------------------------------------------------- 578

Db       2048 GPRTCRNMWSGTFPINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTT 2107

Qy        579 ---------------------RSRRFA----------------------QALPVWARPD 594
                                  |  |||                      ||    ||
Db       2108 DNLKCPCQVPSPEFFTELDGVRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPD 2167

Qy        595 Y-------------------------NPP-----------------------------LV 600
                                        :||                             |:
Db       2168 VAVLTSMLTDPSHITAEAAGRRLARGSPPSVASSSASQLSAPSLKATCTANHDSPDAELI 2227
```

```
Qy        601 ET---WKK-----------------PDYEPPVVHG-----------RSSRRFAQALPV 627
              |    |::                  ::| |               | ||||||||||
Db       2228 EANLLWRQEMGGNITRVESENKVVILDSFDPLVAEEDEREISVPAEILRKSRRFAQALPV 2287

Qy        628 WARPDYNPPLVETWKKPDYEPPVVHG 653
              |||||||||||||||||||||||||
Db       2288 WARPDYNPPLVETWKKPDYEPPVVHG 2313
```

RESULT 14
AAB18541
ID AAB18541 standard; protein; 2955 AA.
XX
AC AAB18541;
XX
DT 15-JAN-2001 (first entry)
XX
DE Polyprotein encoded by sense strand of HCV.
XX
KW Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;
KW viral infectivity; viral replication.
XX
OS Hepatitis C virus.
XX
PN EP1034785-A2.
XX
PD 13-SEP-2000.
XX
PF 16-MAR-1990; 2000EP-00109602.
XX
PR 17-MAR-1989; 89US-00325338.
PR 20-APR-1989; 89US-00341334.
PR 18-MAY-1989; 89US-00355002.
PR 16-MAR-1990; 90EP-00302866.
XX
PA (CHIR) CHIRON CORP.
XX
PI Houghton M, Choo Q, Kuo G;
XX
DR WPI; 2000-566891/53.
DR N-PSDB; AAA75297.
XX
PT Novel composition comprising a hepatitis C virus antisense polynucleotide
PT which is complementary to or corresponds to a sense strand of the virus
PT genome, and selectively hybridizes to it.
XX
PS Example; Fig 17; 75pp; English.
XX
CC The specification describes a pharmaceutical composition which comprises
CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
CC characterized by a positive stranded RNA genome which has 40% homology at
CC the polypeptide level to a HCV polyprotein. The antisense polynucleotide
CC binds to cellular polynucleotides which enhance and/or are required for
CC viral infectivity, replicative ability or chronicity. The antisense
CC polynucleotides may also be designed to bind with high specificity, to be
CC of increased stability, to be stable and to have low toxicity. The
CC composition also comprises an agent which causes viral RNA to be
CC inactive. The composition is used for preventing HCV replication in a
CC system. The present sequence is encoded by a novel HCV cDNA sequence,
CC which is used in the course of the invention
XX
SQ Sequence 2955 AA;

```
  Query Match             36.5%;  Score 1624.5;  DB 3;  Length 2955;
  Best Local Similarity   36.1%;  Pred. No. 9e-107;
  Matches  399;  Conservative   26;  Mismatches   80;  Indels  601;  Gaps   15;

Qy        146 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1211 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 1270

Qy        206 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1271 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 1330

Qy        266 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1331 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 1390

Qy        326 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1391 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 1450

Qy        386 VIDCNTC---------------------------------------------------- 392
              |||||||
Db       1451 VIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPGERPSG 1510

Qy        393 ---------------------------------------------------------- 392

Db       1511 MFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGLTHIDA 1570

Qy        393 ---------------------------------------------------------- 392

Db       1571 HFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLYRLGA 1630

Qy        393 ----------------------------------------------------------ACSG 396
                                                                           ||
Db       1631 VQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRVVLSG 1690

Qy        397 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRGGKPAIVPD 452
              |||||||||||||||||||||||||||||||||||||||||||||    ||  :  | |
Db       1691 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQAE-VIAPA 1749

Qy        453 KEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVTP---DKEI 504
              :  :|: :      |         |:    :      :  |:    | :|    : :
Db       1750 VQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTSPLTTSQTL 1807

Qy        505 LYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGLLG------ 541
              |:                  ||      :   :  |: : ::|     |: |
Db       1808 LFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGVAGALVAFK 1867

Qy        542 ---------------------------------ILRRHVGPGEGAVQWMNRLIAFASRGN 568
                                                  ||||||||||||||||||||||||||
Db       1868 IMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGN 1927

Qy        569 HVSPTHYVPS---------------------------------------------------- 578
              |||||||||
Db       1928 HVSPTHYVPESDAAARVTAILSSLTVTQLLRRLHQWISSECTTPCSGSWLRDIWDWICEV 1987

Qy        579 ---------------------------------------------------------- 578

Db       1988 LSDFKTWLKAKLMPQLPGIPFVSCQRGYKGVWRVDGIMHTRCHCGAEITGHVKNGTMRIV 2047

Qy        579 ---------------------------------------------------------- 578

Db       2048 GPRTCRNMWSGTFPINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTT 2107

Qy        579 ----------------------RSRRFA----------------------QALPVWARPD 594
                                     |  |||                      ||    ||
Db       2108 DNLKCPCQVPSPEFFTELDGVRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPD 2167

Qy        595 Y-------------------------NPP---------------------------LV 600
                                        :||                           |:
Db       2168 VAVLTSMLTDPSHITAEAAGRRLARGSPPSVASSSASQLSAPSLKATCTANHDSPDAELI 2227

Qy        601 ET---WKK-----------------PDYEPPVVHG-----------RSSRRFAQALPV 627
              |    |::                  ::| |               | ||||||||||
Db       2228 EANLLWRQEMGGNITRVESENKVVILDSFDPLVAEEDEREISVPAEILRKSRRFAQALPV 2287

Qy        628 WARPDYNPPLVETWKKPDYEPPVVHG 653
              |||||||||||||||||||||||||
Db       2288 WARPDYNPPLVETWKKPDYEPPVVHG 2313
```

RESULT 15
AAR21519

```
ID   AAR21519 standard; protein; 3011 AA.
XX
AC   AAR21519;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   22-JUN-1992  (first entry)
XX
DE   Compiled HCV sequence.
XX
KW   HCV1; serum; gt11.
XX
OS   Hepatitis C virus type 1.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 9
FT                   /label= ARG
FT   Misc-difference 11
FT                   /label= THR
FT   Misc-difference 176
FT                   /label= THR
FT   Misc-difference 334
FT                   /label= VAL
FT   Misc-difference 603
FT                   /label= ILE
FT   Misc-difference 848
FT                   /label= (ASN)
FT   Misc-difference 1114
FT                   /label= SER
FT   Misc-difference 1117
FT                   /label= THR
FT   Misc-difference 1276
FT                   /label= LEU
FT   Misc-difference 1328
FT                   /label= (VAL)
FT   Misc-difference 1454
FT                   /label= TYR
FT   Misc-difference 1471
FT                   /label= (SER)
FT   Misc-difference 1877
FT                   /label= (GLY)
FT   Misc-difference 1948
FT                   /label= (HIS)
FT   Misc-difference 1949
FT                   /label= (CYS)
FT   Misc-difference 2021
FT                   /label= (VAL)
FT   Misc-difference 2349
FT                   /label= (SER)
FT   Misc-difference 2385
FT                   /label= (PHE)
FT   Misc-difference 2386
FT                   /label= (ALA)
FT   Misc-difference 2502
FT                   /label= (PHE)
FT   Misc-difference 2690
FT                   /label= (GLY)
FT   Misc-difference 2996
FT                   /label= (PRO)
XX
PN   WO9202642-A.
XX
PD   20-FEB-1992.
XX
PF   10-AUG-1990;   90US-00566209.
XX
PR   10-AUG-1990;   90US-00566209.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Houghton M,  Choo QL,  Kuo G,  Weiner AJ,  Uredea MS,  Irvine BD;
PI   Kolberg JA;
XX
DR   WPI; 1992-080094/10.
DR   N-PSDB; AAQ21744.
XX
PT   Reagents for isolating, amplifying and detecting HCV polynucleotide(s) -
PT   used to monitor spread of blood-borne non-a, non-b hepatitis virus
PT   infection and screen blood samples for virus.
XX
PS   Disclosure; Fig 1; 67pp; English.
XX
CC   Heterogeneities in cloned DNAs of HCV1 are indicated by the amino acid
CC   indicated in the features, the parentheses indicated that the
CC   heterogeneity was detected at or near to the 5'- or 3'-end of the HCV in
CC   the clone. The sequence is derived from a composite HCV cDNA from HCV1, a
CC   prototypic HCV. The DNA sequence is based upon sequence information
CC   derived from a no. of HCV cDNA clones, which were isolated form a no. of
CC   HCV cDNA libraries, including the "c" library present in lambda gt11
CC   (ATCC No.40394), and from human serum. The HCV cDNA clones were isolated
CC   by methods described in WO9014436. The clones from which the sequence was
CC   derived are 5'clone32, b114a, 18g, ag30a, CA205a, CA290a, CA216a, pi14a,
CC   CA167b, CA156e, CA84a, CA59a, K9-1 (also called k9-1), 26j, 13i, 12f,
CC   14i, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f,
CC   33f, 33g, 39c, 35f, 19g, 26g, 15e, b5a, 16jh, 6k and p131jh. (Updated on
CC   25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI
CC   field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 3011 AA;

  Query Match             36.5%;  Score 1624.5;  DB 2;  Length 3011;
  Best Local Similarity   36.1%;  Pred. No. 9.2e-107;
  Matches  399;  Conservative   26;  Mismatches   80;  Indels  601;  Gaps   15;

Qy        146 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1211 TDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSK 1270

Qy        206 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1271 AHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIG 1330

Qy        266 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1331 TVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIKGGRH 1390

Qy        326 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1391 LIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYTGDFDS 1450

Qy        386 VIDCNTC------------------------------------------------------ 392
              |||||||
Db       1451 VIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGINRFVAPGERPSG 1510

Qy        393 ------------------------------------------------------------ 392

Db       1511 MFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGLTHIDA 1570

Qy        393 ------------------------------------------------------------ 392

Db       1571 HFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLYRLGA 1630

Qy        393 ------------------------------------------------------------ACSG 396
                                                                           ||
Db       1631 VQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRVVLSG 1690

Qy        397 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGL----SRGGKPAIVPD 452
              |||||||||||||||||||||||||||||||||||||||||||||    || :  | |
Db       1691 KPAIIPDREVLYREFDEMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQAE-VIAPA 1749

Qy        453 KEVLYQQYD-----EMEECSQAAPYIEQAQVIAHQFKEKVLGLIDNDQVVVTP---DKEI 504
               :  :|: :      |         |:     :       :  |:     | :|    : :
Db       1750 VQTNWQKLETFWAKHMWNFISGIQYLAGLSTLPG--NPAIASLMAFTAAVTSPLTTSQTL 1807

Qy        505 LYE-----------------AFDEMEECASKAALIEEGQRMAEMLKSKIQGLLG------ 541
              |:                 ||       :    :  |: : ::|     |: |
```

```
Db       1808 LFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLIDILAGYGAGVAGALVAFK 1867

Qy        542 --------------------------------ILRRHVGPGEGAVQWMNRLIAFASRGN 568
                                               ||||||||||||||||||||||||||
Db       1868 IMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGN 1927

Qy        569 HVSPTHYVPS---------------------------------------------------- 578
              |||||||||
Db       1928 HVSPTHYVPESDAAARVTAILSSLTVTQLLRRLHQWISSECTTPCSGSWLRDIWDWICEV 1987

Qy        579 ---------------------------------------------------------- 578

Db       1988 LSDFKTWLKAKLMPQLPGIPFVSCQRGYKGVWRVDGIMHTRCHCGAEITGHVKNGTMRIV 2047

Qy        579 ---------------------------------------------------------- 578

Db       2048 GPRTCRNMWSGTFPINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTT 2107

Qy        579 ---------------------RSRRFA-----------------------QALPVWARPD 594
                                   |  |||                       ||    ||
Db       2108 DNLKCPCQVPSPEFFTELDGVRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPD 2167

Qy        595 Y-------------------------NPP-----------------------------LV 600
                                       :||                             |:
Db       2168 VAVLTSMLTDPSHITAEAAGRRLARGSPPSVASSSASQLSAPSLKATCTANHDSPDAELI 2227

Qy        601 ET---WKK-----------------PDYEPPVVHG-----------RSSRRFAQALPV 627
              |    |::                    ::| |               | ||||||||||
Db       2228 EANLLWRQEMGGNITRVESENKVVILDSFDPLVAEEDEREISVPAEILRKSRRFAQALPV 2287

Qy        628 WARPDYNPPLVETWKKPDYEPPVVHG 653
              |||||||||||||||||||||||||
Db       2288 WARPDYNPPLVETWKKPDYEPPVVHG 2313
```

Search completed: June 21, 2004, 10:30:18
Job time : 64.4736 secs

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:18:09 ; Search time 40.9396 Seconds
(without alignments)
4734.482 Million cell updates/sec

Title: US-10-658-782-2
Perfect score: 3619
Sequence: 1 MAPITAYAQQTRGLLGCIIT.....PAIIPDREVLYREFDEMEEC 686

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3619	100.0	686	5	AAU76377	Aau76377 Hepatitis
2	3619	100.0	686	5	AAE18689	Aae18689 HCV-1 NS3
3	3619	100.0	686	6	ABG72261	Abg72261 HCV-1 NS3
4	3619	100.0	686	7	ADC06767	Adc06767 HCV mutan
5	3602	99.5	2261	1	AAP90164	Aap90164 Peptide e
6	3602	99.5	2436	1	AAP92050	Aap92050 Sequence
7	3602	99.5	2436	1	AAP90288	Aap90288 Peptide e
8	3602	99.5	2772	3	AAB18540	Aab18540 Protein e
9	3602	99.5	2955	2	AAY14975	Aay14975 Amino aci
10	3602	99.5	2955	3	AAB18541	Aab18541 Polyprote
11	3602	99.5	3011	2	AAR90931	Aar90931 Hepatitis
12	3602	99.5	3011	2	AAW34480	Aaw34480 HCV polyp
13	3602	99.5	3011	2	AAW40038	Aaw40038 HCV polyp
14	3602	99.5	3011	5	AAE22049	Aae22049 Hepatitis
15	3600	99.5	728	5	AAE18688	Aae18688 NS3/4a mu
16	3600	99.5	728	7	ADC06766	Adc06766 HCV mutan
17	3599	99.4	2301	1	AAP92047	Aap92047 Sequence
18	3595	99.3	2772	2	AAR08123	Aar08123 Hepatitis
19	3594	99.3	686	4	AAB62633	Aab62633 HCV NS34A
20	3593	99.3	3011	2	AAR21519	Aar21519 Compiled
21	3590	99.2	2435	2	AAR25135	Aar25135 HCV polyp
22	3589	99.2	3011	2	AAR31621	Aar31621 Hepatitis
23	3587	99.1	3011	5	AAU84597	Aau84597 HCV polyp
24	3586	99.1	2816	2	AAR34009	Aar34009 HCV-1 pol
25	3583	99.0	1786	1	AAP90158	Aap90158 Protein s
26	3583	99.0	2436	2	AAR28582	Aar28582 HCV amino
27	3583	99.0	2894	2	AAR70230	Aar70230 Composite
28	3580	98.9	2894	2	AAR24440	Aar24440 Composite
29	3579	98.9	1766	1	AAP92041	Aap92041 Hepatitis
30	3565	98.5	686	5	AAE21837	Aae21837 Hepatitis
31	3565	98.5	686	5	AAE19900	Aae19900 Hepatitis
32	3565	98.5	686	7	ABW00351	Abw00351 Hepatitis
33	3561	98.4	686	5	AAE21838	Aae21838 Hepatitis
34	3561	98.4	686	5	AAE19907	Aae19907 Hepatitis
35	3561	98.4	686	7	ABW00358	Abw00358 Hepatitis
36	3561	98.4	3011	2	AAR40120	Aar40120 HCV genom
37	3560	98.4	686	5	AAE21840	Aae21840 Hepatitis
38	3560	98.4	686	5	AAE19919	Aae19919 Hepatitis
39	3560	98.4	686	7	ABW00370	Abw00370 Hepatitis
40	3560	98.4	2955	2	AAR08124	Aar08124 Hepatitis
41	3560	98.4	3011	2	AAW77397	Aaw77397 Hepatitis
42	3560	98.4	3011	6	ABP71460	Abp71460 Amino aci
43	3560	98.4	3012	5	AAU99289	Aau99289 Hepatitis
44	3560	98.4	3012	6	ABU61848	Abu61848 HCV H77 c
45	3559	98.3	686	5	AAE21843	Aae21843 Hepatitis

ALIGNMENTS

RESULT 1
AAU76377
ID AAU76377 standard; protein; 686 AA.
XX
AC AAU76377;
XX
DT 08-MAY-2002 (first entry)
XX
DE Hepatitis C virus NS3/4a conformational epitope protein sequence.
XX
KW Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion;
KW immunoassay solid support; multiple epitope fusion antigen; MEFA;
KW non-structural protein; mutant; mutein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 403
FT /note= "Wild-type Thr substituted by Pro"
FT Misc-difference 404
FT /note= "Wild-type Ser substituted by Ile"
XX
PN WO200196870-A2.
XX
PD 20-DEC-2001.
XX
PF 14-JUN-2001; 2001WO-US019156.
XX
PR 15-JUN-2000; 2000US-0212082P.
PR 02-APR-2001; 2001US-0280811P.
PR 02-APR-2001; 2001US-0280867P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Chien DY, Arcangel P, Tandeske L, George-Nasciemento C, Coit D;
PI Medina-Selby A;
XX
DR WPI; 2002-090228/12.
DR N-PSDB; ABK15344.
XX
PT Immunoassay solid support, useful for detecting hepatitis C virus
PT infection in biological sample, comprises HCV NS3/4a conformational
PT epitope and multiple epitope fusion antigen bound to the support.
XX
PS Claim 5; Fig 3; 92pp; English.
XX
CC The present invention relates to a new immunoassay solid support

CC consisting essentially of at least one hepatitis C virus (HCV) NS3/4a
CC conformational epitope and a multiple epitope fusion antigen (MEFA),
CC bound to the support. The NS3/4a conformational epitope and/or MEFA
CC reacts specifically with anti-HCV antibodies present in a biological
CC sample from an HCV-infected individual. The immunoassay of the invention
CC is useful for detecting hepatitis C virus infection in a biological
CC sample. The method of the invention provides a sensitive, accurate
CC diagnostic and prognostic tool to provide adequate patient care and to
CC prevent transmission of HCV by blood and by blood products, or by
CC personal contact. Use of NS3/4a conformational epitope in combination
CC with MEFA, provides a sensitive and reliable method for detecting early
CC HCV seroconversion. Use of MEFA has the added advantages of decreasing
CC masking problems, improving sensitivity in detecting antibodies by
CC allowing a greater number of epitopes on a unit surface area of
CC substrate, and improving substrate. Detection accuracy is increased and
CC the incidence of false results is reduced because of the identification
CC and the use of highly immunogenic HCV antigens which are present during
CC the early stages of HCV seroconversion. The present amino acid sequence
CC represents the non-structural protein NS3/4a conformational epitope of
CC the invention
XX
SQ Sequence 686 AA;

Query Match 100.0%; Score 3619; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 1.6e-306;
Matches 686; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
              ||||||||||||||||||||||||||
Db        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
```

RESULT 2
AAE18689
ID AAE18689 standard; protein; 686 AA.
XX
AC AAE18689;
XX
DT 17-MAY-2002 (first entry)
XX
DE HCV-1 NS3/4a mutant conformational antigen.
XX
KW Hepatitis C virus; NS3/4a antigen; HCV infection; mutant; mutein.
XX
OS Hepatitis C virus type 1.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 403
FT /note= "Wild type Thr substituted with Pro"
FT Misc-difference 404
FT /note= "Wild type Ser substituted with Ile"
XX
PN WO200196875-A2.
XX
PD 20-DEC-2001.
XX
PF 14-JUN-2001; 2001WO-US019369.
XX
PR 15-JUN-2000; 2000US-0212082P.
PR 02-APR-2001; 2001US-0280811P.
PR 02-APR-2001; 2001US-0280867P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Chien DY, Arcangel P, Tandeske L, George-Nasciemento C, Coit D;
PI Medina-Selby A;
XX
DR WPI; 2002-179522/23.
DR N-PSDB; AAD29795.
XX
PT Immunoassay solid support useful for detecting hepatitis C virus
PT infection in a biological sample, comprises at least one of HCV anti-core
PT antibody and HCV NS3/4a epitope, bound to the support.
XX
PS Example 2; Fig 4; 87pp; English.
XX
CC The present invention relates to hepatitis C virus (HCV) core antigen and
CC NS (nonstructural) 3/4a antibody combination assay that can detect both
CC HCV antigens and antibodies present in a sample using a single solid
CC matrix as well as immunoassay solid supports for use in the assay. The
CC solid support is useful for detecting HCV infection in a biological
CC sample. The present sequence is HCV-1 NS3/4a mutant conformational
CC antigen. This sequence is used in the exemplification of the invention
XX
SQ Sequence 686 AA;

Query Match 100.0%; Score 3619; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 1.6e-306;
Matches 686; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
```

```
Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
              |||||||||||||||||||||||||
Db        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
```

RESULT 3
ABG72261
ID ABG72261 standard; protein; 686 AA.
XX
AC ABG72261;
XX
DT 06-MAR-2003 (first entry)
XX
DE HCV-1 NS3/4a conformational antigen.
XX
KW Immunoassay solid support; Hepatitis C Virus type-1; HCV-1;
KW NS3/4a conformational epitope; multiple epitope fusion antigen; MEFA;
KW anti-HCV antibody; NS3/4a conformational antigen; HCV infection; mutant;
KW mutein.
XX
OS Hepatitis C virus type 1.
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	Region	2. .686
FT		/note= "Corresponds to amino acid residues 1027-1711 of
FT		HCV-1 NS3/4a polypeptide"
FT	Misc-difference	403
FT		/note= "Substitution of wild-type Thr to Pro"
FT	Misc-difference	404
FT		/note= "Substitution of wild-type Ser to Ile"

XX
PN US2002146685-A1.
XX
PD 10-OCT-2002.
XX
PF 14-JUN-2001; 2001US-00881654.
XX
PR 15-JUN-2000; 2000US-0212082P.
PR 02-APR-2001; 2001US-0280811P.
PR 02-APR-2001; 2001US-0280867P.
XX
PA (CHIE/) CHIEN D Y.
PA (ARCA/) ARCANGEL P.
PA (TAND/) TANDESKE L.
PA (GEOR/) GEORGE-NASCIMENTO C.
PA (COIT/) COIT D.
PA (MEDI/) MEDINA-SELBY A.
XX
PI Chien DY, Arcangel P, Tandeske L, George-Nascimento C, Coit D;
PI Medina-Selby A;
XX
DR WPI; 2003-147573/14.
DR N-PSDB; ABX14410.
XX
PT Immunoassay solid support for detecting Hepatitis C Virus infection in
PT biological samples, comprises Hepatitis C Virus conformational epitope
PT and multiple epitope fusion antigen.
XX
PS Claim 2; Fig 3A-3D; 45pp; English.
XX
CC The present invention relates to immunoassays comprising Hepatitis C
CC Virus (HCV) NS3/4a conformational epitope and multiple epitope fusion
CC antigen (MEFA), bound to a solid support. The NS3/4a epitope and/or the
CC multiple epitope fusion antigen react with anti-HCV antibodies present in
CC a biological sample from an HCV-infected individual. The immunoassays and
CC methods of the invention are useful for detecting HCV infection in a
CC biological sample. The inventive immunoassay solid support provides a
CC sensitive and reliable method for detecting early HCV seroconversion. The
CC assays can detect HCV infection caused by any six known genotypes of HCV.
CC The use of the multiple epitope fusion proteins decreases masking
CC problems, improves sensitivity in detecting antibodies by allowing a
CC greater number of epitopes on a unit area of substrate, and improves
CC selectivity. The present sequence represents HCV type 1 (HCV-1) NS3/4a
CC conformational antigen, a mutant of the HCV-1 NS3/4a polypeptide
XX
SQ Sequence 686 AA;

Query Match 100.0%; Score 3619; DB 6; Length 686;
Best Local Similarity 100.0%; Pred. No. 1.6e-306;
Matches 686; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
```

```
Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
              |||||||||||||||||||||||||
Db        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
```

RESULT 4
ADC06767
ID ADC06767 standard; protein; 686 AA.
XX
AC ADC06767;
XX
DT 18-DEC-2003 (first entry)
XX
DE HCV mutant conformational NS3/4a epitope protein T403P/S404I.
XX
KW immunoassay solid support; HCV; NS3/4a; non-structural;
KW non-A, non-B hepatitis; NANB; conformational epitope; mutant; mutein.
XX
OS Synthetic.
OS Hepatitis C virus.
XX
FH Key   Location/Qualifiers
FT Misc-difference 303
FT   /note= "Wild-type Thr replaced by Pro"
FT Misc-difference 304
FT   /note= "Wild-type Ser replaced by Ile"
XX
PN US2002192639-A1.
XX
PD 19-DEC-2002.
XX
PF 14-JUN-2001; 2001US-00881239.
XX
PR 15-JUN-2000; 2000US-0212082P.
PR 02-APR-2001; 2001US-0280811P.
PR 02-APR-2001; 2001US-0280867P.
XX
PA (CHIE/) CHIEN D Y.
PA (ARCA/) ARCANGEL P.
PA (TAND/) TANDESKE L.
PA (GEOR/) GEORGE-NASCIMENTO C.
PA (COIT/) COIT D.
PA (MEDI/) MEDINA-SELBY A.
XX
PI Chien DY, Arcangel P, Tandeske L, George-Nascimento C, Coit D;
PI Medina-Selby A;
XX
DR WPI; 2003-644609/61.
DR N-PSDB; ADC06768.
XX
PT Immunoassay solid support for detecting hepatitis C virus infection in
PT biological samples, comprises a hepatitis C virus anti-core antibody and
PT an isolated hepatitis C virus NS3/4a epitope bound HCV anti-core
PT antibody.
XX
PS Claim 6; Fig 4; 40pp; English.
XX
CC The invention relates to a novel immunoassay solid support comprising at
CC least one hepatitis C virus (HCV) anti-core antibody and at least one
CC isolated HCV NS3/4a (non-structural protein 3/4a) epitope bound thereto.
CC The system of the invention may be useful for detecting HCV infection in
CC a biological sample and for treating or detecting non-A, non-B hepatitis
CC (NANB hepatitis). The current sequence is that of the HCV mutant
CC conformational NS3/4a epitope protein of the invention which contains
CC T403P/S404I mutations.
XX
SQ Sequence 686 AA;

```
  Query Match             100.0%;  Score 3619;  DB 7;  Length 686;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-306;
  Matches  686;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
              |||||||||||||||||||||||||
Db        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
```

RESULT 5
AAP90164
ID AAP90164 standard; protein; 2261 AA.
XX
AC AAP90164;
XX
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Peptide encoded by composite hepatitis C virus cDNA.
XX
KW Hepatitis C virus; clone 12f; clone 15e; probe; vaccine.

XX
OS Pan troglodytes.
XX
PN GB2212511-A.
XX
PD 26-JUL-1989.
XX
PF 18-NOV-1988; 88GB-00027024.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 26-OCT-1988; 88US-00263584.
XX
PA (CHIR) CHIRON CORP.
XX
PI Houghton M, Choo QL, Kuo G;
XX
DR WPI; 1989-215054/30.
DR N-PSDB; AAN90331.
XX
PT Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
PS Disclosure; Fig 32; 30pp; English.
XX
CC The sequence is the peptide encoded by the composite hepatitis C virus
CC (HCV) cDNA of AAN90331. The polypeptides are used to diagnose HCV-induced
CC NANBH, to raise antibodies for immunoassay or treatment, or to produce
CC vaccines. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 2261 AA;

Query Match 99.5%; Score 3602; DB 1; Length 2261;
Best Local Similarity 99.6%; Pred. No. 2.8e-304;
Matches 683; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 460

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 520

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 580

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 640

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 700

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 760

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYT 820

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 880

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        881 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 940

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 1000

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1001 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 1060

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
              |||||||||||||||||||||||||
Db       1061 VVLSGKPAIIPDREVLYREFDEMEEC 1086
```

RESULT 6
AAP92050
ID AAP92050 standard; protein; 2436 AA.
XX
AC AAP92050;
XX
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX
DE Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones K9
DE -1 through 15e.
XX
KW Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH).
XX
OS Hepatitis C virus.
XX
PN EP318216-A.
XX
PD 31-MAY-1989.
XX
PF 18-NOV-1988; 88EP-00310922.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 06-MAY-1988; 88US-00191263.
PR 26-OCT-1988; 88US-00263584.
PR 14-NOV-1988; 88US-00271450.
XX
PA (CHIR) CHIRON CORP.
PA (CHIR) CHIRON CORP.
XX
PI Houghton M, Choo QL, Kuo G;
XX
DR WPI; 1989-159274/22.
DR N-PSDB; AAN92106.
XX
PT Purified hepatitis C virus - and associated nucleic acids and
PT polypeptide(s).
XX
PS Claim 13; Fig 47-1-47-8; 139pp; English.
XX
CC It is the sequence encoded in the open reading frame of hepatitis C virus
CC (HCV) cDNA inserts in clones K9-1 through 15e. It is antigenic and could
CC be used in immunoassay reagents and vaccines and to generate antibodies
CC useful in diagnosis and passive immunotherapy for HCV infection/non-A,
CC non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.) (Updated
CC on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 2436 AA;

Query Match 99.5%; Score 3602; DB 1; Length 2436;
Best Local Similarity 99.6%; Pred. No. 3.1e-304;
Matches 683; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 635
```

```
Qy          61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         636 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 695

Qy         121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         696 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 755

Qy         181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         756 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 815

Qy         241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         816 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 875

Qy         301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         876 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 935

Qy         361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
               |||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db         936 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYT 995

Qy         421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         996 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 1055

Qy         481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1056 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 1115

Qy         541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1116 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 1175

Qy         601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1176 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 1235

Qy         661 VVLSGKPAIIPDREVLYREFDEMEEC 686
               |||||||||||||||||||||||||
Db        1236 VVLSGKPAIIPDREVLYREFDEMEEC 1261
```

RESULT 7
AAP90288
ID AAP90288 standard; protein; 2436 AA.
XX
AC AAP90288;
XX
DT 25-MAR-2003 (revised)
DT 19-JUL-2001 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Peptide encoded by composite hepatitis C cDNA.
XX
KW Hepatitis C virus; clone 15e; clone k9-1; probe; vaccine.
XX
OS Pan troglodytes.
XX
PN GB2212511-A.
XX
PD 26-JUL-1989.
XX
PF 18-NOV-1988; 88GB-00027024.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 26-OCT-1988; 88US-00263584.
XX
PA (CHIR) CHIRON CORP.

XX
PI Houghton M, Choo QL, Kuo G;
XX
DR WPI; 1989-215054/30.
DR N-PSDB; AAN90336.
XX
PT Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
PS Disclosure; Fig 47-1 to 47-8; 30pp; English.
XX
CC The sequence is the peptide encoded by the composite hepatitis C virus
CC (HCV) cDNA of AAN90336. The polypeptides are used to diagnose HCV-induced
CC NANBH, to raise antibodies for immunoassay or treatment, or to produce
CC vaccines. (N.B. This record was resubmitted to correct errors in the
CC sequence.) (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 2436 AA;

Query Match 99.5%; Score 3602; DB 1; Length 2436;
Best Local Similarity 99.6%; Pred. No. 3.1e-304;
Matches 683; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy           1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
               :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         576 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 635

Qy          61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         636 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 695

Qy         121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         696 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 755

Qy         181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         756 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 815

Qy         241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         816 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 875

Qy         301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         876 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 935

Qy         361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
               |||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db         936 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYT 995

Qy         421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         996 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 1055

Qy         481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1056 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 1115

Qy         541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1116 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 1175

Qy         601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1176 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 1235

Qy         661 VVLSGKPAIIPDREVLYREFDEMEEC 686
               |||||||||||||||||||||||||
Db        1236 VVLSGKPAIIPDREVLYREFDEMEEC 1261
```

```
RESULT 8
AAB18540
ID   AAB18540 standard; protein; 2772 AA.
XX
AC   AAB18540;
XX
DT   15-JAN-2001  (first entry)
XX
DE   Protein encoded by a cDNA compiled Hepatitis C virus cDNA clones.
XX
KW   Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;
KW   viral infectivity; viral replication.
XX
OS   Hepatitis C virus.
XX
PN   EP1034785-A2.
XX
PD   13-SEP-2000.
XX
PF   16-MAR-1990; 2000EP-00109602.
XX
PR   17-MAR-1989;   89US-00325338.
PR   20-APR-1989;   89US-00341334.
PR   18-MAY-1989;   89US-00355002.
PR   16-MAR-1990;   90EP-00302866.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Houghton M,  Choo Q,  Kuo G;
XX
DR   WPI; 2000-566891/53.
DR   N-PSDB; AAA75296.
XX
PT   Novel composition comprising a hepatitis C virus antisense polynucleotide
PT   which is complementary to or corresponds to a sense strand of the virus
PT   genome, and selectively hybridizes to it.
XX
PS   Example; Fig 16; 75pp; English.
XX
CC   The specification describes a pharmaceutical composition which comprises
CC   a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
CC   characterized by a positive stranded RNA genome which has 40% homology at
CC   the polypeptide level to a HCV polyprotein. The antisense polynucleotide
CC   binds to cellular polynucleotides which enhance and/or are required for
CC   viral infectivity, replicative ability or chronicity. The antisense
CC   polynucleotides may also be designed to bind with high specificity, to be
CC   of increased stability, to be stable and to have low toxicity. The
CC   composition also comprises an agent which causes viral RNA to be
CC   inactive. The composition is used for preventing HCV replication in a
CC   system. The present sequence is encoded by a novel HCV cDNA sequence,
CC   which is used in the course of the invention
XX
SQ   Sequence 2772 AA;

  Query Match             99.5%;  Score 3602;  DB 3;  Length 2772;
  Best Local Similarity   99.6%;  Pred. No. 3.8e-304;
  Matches  683;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 971

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 1031

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1032 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 1091

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1092 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 1151

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1152 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 1211

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1212 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 1271

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db       1272 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYT 1331

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1332 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 1391

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1392 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 1451

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1452 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 1511

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1512 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 1571

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
              |||||||||||||||||||||||||
Db       1572 VVLSGKPAIIPDREVLYREFDEMEEC 1597


RESULT 9
AAY14975
ID   AAY14975 standard; protein; 2955 AA.
XX
AC   AAY14975;
XX
DT   20-MAR-2003  (revised)
DT   08-NOV-1999  (first entry)
XX
DE   Amino acid sequence of HCV-1 ORF.
XX
KW   Hepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH;
KW   HCV infection; vaccine.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 441
FT                   /note= "encoded by TT"
FT   Misc-difference 461
FT                   /note= "encoded by CCCC"
XX
PN   EP939128-A2.
XX
PD   01-SEP-1999.
XX
PF   17-SEP-1990;   99EP-00101746.
XX
PR   15-SEP-1989;   89US-00408045.
PR   21-DEC-1989;   89US-00456142.
PR   17-SEP-1990;   90EP-00310149.
XX
PA   (OYAA/) OYA A.
PA   (CHIR ) CHIRON CORP.
XX
PI   Miyamura T,  Saito I,  Houghton M,  Weiner AJ,  Han J,  Kolberg JA;
PI   Cha T,  Irvine BD;
XX
DR   WPI; 1999-480843/41.
```

```
DR   N-PSDB; AAZ07656.
XX
PT   New Hepatitis C Virus isolates, useful for diagnosis of hepatitis
PT   infections and development of vaccines.
XX
PS   Disclosure; Fig 12; 132pp; English.
XX
CC   The invention provides two new isolates of hepatitis C virus (HCV), J1
CC   and J7. These two isolates comprise nucleotide and amino acid sequences
CC   that are distinct from the HCV isolate HCV-1. The nucleotide sequences
CC   may be used to detect non-A, non-B HCV (NANBH) polynucleotides by
CC   hybridisation for diagnosis of NANBH infections. They may also be used to
CC   screen blood donors, donated blood and blood products for this infection.
CC   The isolates may also be used to isolate other naturally occurring
CC   variants of the virus. The polypeptides may be used as a vaccine for
CC   administration to patients to protect against infection with NANBH. The
CC   present sequence represents the amino acid sequence of HCV-1 ORF.
CC   (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
CC   correct PR field.)
XX
SQ   Sequence 2955 AA;

  Query Match             99.5%;  Score 3602;  DB 2;  Length 2955;
  Best Local Similarity   99.6%;  Pred. No. 4.2e-304;
  Matches  683;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 1085

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 1145

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 1205

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 1265

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 1325

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 1385

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
              |||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db       1386 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYT 1445

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 1505

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 1565

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 1625

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 1685

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
              |||||||||||||||||||||||||
Db       1686 VVLSGKPAIIPDREVLYREFDEMEEC 1711
```

```
RESULT 10
AAB18541
ID   AAB18541 standard; protein; 2955 AA.
XX
AC   AAB18541;
XX
DT   15-JAN-2001  (first entry)
XX
DE   Polyprotein encoded by sense strand of HCV.
XX
KW   Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;
KW   viral infectivity; viral replication.
XX
OS   Hepatitis C virus.
XX
PN   EP1034785-A2.
XX
PD   13-SEP-2000.
XX
PF   16-MAR-1990; 2000EP-00109602.
XX
PR   17-MAR-1989;   89US-00325338.
PR   20-APR-1989;   89US-00341334.
PR   18-MAY-1989;   89US-00355002.
PR   16-MAR-1990;   90EP-00302866.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Houghton M,  Choo Q,  Kuo G;
XX
DR   WPI; 2000-566891/53.
DR   N-PSDB; AAA75297.
XX
PT   Novel composition comprising a hepatitis C virus antisense polynucleotide
PT   which is complementary to or corresponds to a sense strand of the virus
PT   genome, and selectively hybridizes to it.
XX
PS   Example; Fig 17; 75pp; English.
XX
CC   The specification describes a pharmaceutical composition which comprises
CC   a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
CC   characterized by a positive stranded RNA genome which has 40% homology at
CC   the polypeptide level to a HCV polyprotein. The antisense polynucleotide
CC   binds to cellular polynucleotides which enhance and/or are required for
CC   viral infectivity, replicative ability or chronicity. The antisense
CC   polynucleotides may also be designed to bind with high specificity, to be
CC   of increased stability, to be stable and to have low toxicity. The
CC   composition also comprises an agent which causes viral RNA to be
CC   inactive. The composition is used for preventing HCV replication in a
CC   system. The present sequence is encoded by a novel HCV cDNA sequence,
CC   which is used in the course of the invention
XX
SQ   Sequence 2955 AA;

  Query Match             99.5%;  Score 3602;  DB 3;  Length 2955;
  Best Local Similarity   99.6%;  Pred. No. 4.2e-304;
  Matches  683;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 1085

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 1145

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 1205

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
```

```
Db       1206 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 1265

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300

Db       1266 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 1325

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360

Db       1326 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 1385

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420

Db       1386 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYT 1445

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480

Db       1446 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 1505

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540

Db       1506 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 1565

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600

Db       1566 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 1625

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660

Db       1626 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 1685

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686

Db       1686 VVLSGKPAIIPDREVLYREFDEMEEC 1711
```

```
RESULT 11
AAR90931
ID   AAR90931 standard; protein; 3011 AA.
XX
AC   AAR90931;
XX
DT   25-MAR-2003  (revised)
DT   15-MAY-1996  (first entry)
XX
DE   Hepatitis C virus polyprotein.
XX
KW   Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;
KW   antibodies.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .122
FT                   /label= antigen
FT                   /note= "C22; AAR90936"
FT   Misc-difference 199. .328
FT                   /label= antigen
FT                   /note= "S2; AAR90935"
FT   Misc-difference 1192. .1457
FT                   /label= antigen
FT                   /note= "C33c; AAR90932"
FT   Misc-difference 1569. .1931
FT                   /label= antigen
FT                   /note= "C100; AAR90933"
FT   Misc-difference 2054. .2464
FT                   /label= antigen
FT                   /note= "NS5; AAR90934"
XX
PN   EP693687-A1.
XX
PD   24-JAN-1996.
XX
PF   03-APR-1991;   95EP-00114016.
XX
PR   04-APR-1990;   90US-00504352.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Houghton M,  Choo Q,  Kuo G;
XX
DR   WPI; 1996-117956/13.
DR   N-PSDB; AAT12710.
XX
PT   Combinations of synthetic Hepatitis C Virus antigens - provide more
PT   effective diagnosis of Non-A, Non-B Hepatitis.
XX
PS   Disclosure; Fig 1(A-Y); 53pp; English.
XX
CC   The combination comprises an HCV antigen from the C domain (pref. C22 -
CC   AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -
CC   AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5
CC   (AAR90934) domain. The antigens may in the form of a fusion protein, a
CC   simple physical mixture, or the individual antigens commonly bound to a
CC   solid matrix. They are pref. prepd. by recombinant DNA techniques
CC   (primers are given in AAT12711-T12716), but can be synthesised or
CC   isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003
CC   to correct PF field.)
XX
SQ   Sequence 3011 AA;

  Query Match             99.5%;  Score 3602;  DB 2;  Length 3011;
  Best Local Similarity   99.6%;  Pred. No. 4.3e-304;
  Matches  683;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60

Db       1026 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 1085

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120

Db       1086 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 1145

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180

Db       1146 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 1205

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240

Db       1206 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 1265

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300

Db       1266 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 1325

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360

Db       1326 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 1385

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420

Db       1386 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYT 1445

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480

Db       1446 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 1505

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540

Db       1506 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 1565

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600

Db       1566 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 1625

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 1685

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
              |||||||||||||||||||||||||
Db       1686 VVLSGKPAIIPDREVLYREFDEMEEC 1711


RESULT 12
AAW34480
ID   AAW34480 standard; protein; 3011 AA.
XX
AC   AAW34480;
XX
DT   25-MAR-2003  (revised)
DT   16-MAR-1998  (first entry)
XX
DE   HCV polyprotein.
XX
KW   PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
KW   C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
KW   NS4.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 366
FT                   /note= "can optionally be Arg"
FT   Misc-difference 372
FT                   /note= "can optionally be Thr"
FT   Misc-difference 867
FT                   /note= "can optionally be Thr"
FT   Misc-difference 1341
FT                   /note= "can optionally be Val"
FT   Misc-difference 2148
FT                   /note= "can optionally be Ile"
FT   Misc-difference 2883
FT                   /note= "can optionally be Asn"
FT   Misc-difference 3681
FT                   /note= "can optionally be Ser"
FT   Misc-difference 3690
FT                   /note= "can optionally be Thr"
FT   Misc-difference 4167
FT                   /note= "can optionally be Leu"
FT   Misc-difference 4323
FT                   /note= "can optionally be Val"
FT   Misc-difference 4701
FT                   /note= "can optionally be Tyr"
FT   Misc-difference 4752
FT                   /note= "can optionally be Ser"
FT   Misc-difference 5970
FT                   /note= "can optionally be Gly"
FT   Misc-difference 6183
FT                   /note= "can optionally be His"
FT   Misc-difference 6186
FT                   /note= "can optionally be Cys"
FT   Misc-difference 6402
FT                   /note= "can optionally be Val"
FT   Misc-difference 7386
FT                   /note= "can optionally be Ser"
FT   Misc-difference 7494
FT                   /note= "can optionally be Phe"
FT   Misc-difference 7497
FT                   /note= "can optionally be Ala"
FT   Misc-difference 7845
FT                   /note= "can optionally be Phe"
FT   Misc-difference 8409
FT                   /note= "can optionally be Gly"
FT   Misc-difference 9102
FT                   /note= "can optionally be Gly"
FT   Misc-difference 9327
FT                   /note= "can optionally be Pro"
XX
PN   US5683864-A.
XX
PD   04-NOV-1997.
XX
PF   07-JUL-1992;   92US-00910760.
XX
PR   18-NOV-1987;   87US-00122714.
PR   30-DEC-1987;   87US-00139886.
PR   26-FEB-1988;   88US-00161072.
PR   06-MAY-1988;   88US-00191263.
PR   26-OCT-1988;   88US-00263584.
PR   14-NOV-1988;   88US-00271450.
PR   17-MAR-1989;   89US-00325338.
PR   20-APR-1989;   89US-00341334.
PR   21-APR-1989;   89US-00353896.
PR   18-MAY-1989;   89US-00355002.
PR   04-APR-1990;   90US-00504352.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kuo G,  Houghton M,  Choo Q;
XX
DR   WPI; 1997-548976/50.
DR   N-PSDB; AAT99981.
XX
PT   Combination of three hepatitis C virus antigens - used for detection of
PT   specific antibodies to diagnose infection.
XX
PS   Disclosure; Col 25-46; 57pp; English.
XX
CC   This sequence represents the Hepatitis C virus polyprotein. Fragments of
CC   the DNA encoding this sequence can be amplified and used in the
CC   combination of HCV antigens of the invention. The HCV antigen combination
CC   comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)
CC   1-120 of the HCV polyprotein), or its immunologically reactive fragment
CC   containing at least 8 aa. It also comprises two additional antigens from
CC   two different polyprotein domains, including at least 8 aa from the NS3,
CC   NS4, S or NS5 domains of the polyprotein, corresponding, respectively, to
CC   aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polyprotein.
CC   Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa
CC   regions of the HCV polyprotein. These antigen combinations are used
CC   diagnostically to detect anti-HCV antibodies, using any standard
CC   immunoassay format. These antigen combinations have a broader range of
CC   reactivity with antibodies than any antigen individually. (Updated on 25-
CC   MAR-2003 to correct PR field.)
XX
SQ   Sequence 3011 AA;

  Query Match             99.5%;  Score 3602;  DB 2;  Length 3011;
  Best Local Similarity   99.6%;  Pred. No. 4.3e-304;
  Matches  683;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 1085

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 1145

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 1205

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 1265

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 1325

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 1385

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
                ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db       1386 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYT 1445

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 1505

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 1565

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 1625

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 1685

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
                |||||||||||||||||||||||||
Db       1686 VVLSGKPAIIPDREVLYREFDEMEEC 1711


RESULT 13
AAW40038
ID   AAW40038 standard; protein; 3011 AA.
XX
AC   AAW40038;
XX
DT   26-MAY-1998  (first entry)
XX
DE   HCV polyprotein.
XX
KW   Hepatitis C virus C domain; HCV; C antigen; immunological activity;
KW   NS3 domain; NS4 domain; S domain; NS5 domain.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .120
FT                   /label= C_domain
FT   Modified-site   9
FT                   /note= "As given in the specification this amino acid can
FT                   also be Arg"
FT   Modified-site   11
FT                   /note= "As given in the specification this amino acid can
FT                   also be Thr"
FT   Domain          120. .400
FT                   /label= S_domain
FT   Modified-site   174
FT                   /note= "As given in the specification this amino acid can
FT                   also be Thr"
FT   Modified-site   334
FT                   /note= "As given in the specification this amino acid can
FT                   also be Val"
FT   Modified-site   603
FT                   /note= "As given in the specification this amino acid can
FT                   also be Ile"
FT   Modified-site   847
FT                   /note= "As given in the specification this amino acid can
FT                   also be Asn"
FT   Domain          1050. .1640
FT                   /label= NS3_domain
FT   Modified-site   1114
FT                   /note= "As given in the specification this amino acid can
FT                   also be Ser"
FT   Modified-site   1217
FT                   /note= "As given in the specification this amino acid can
FT                   also be Thr"
FT   Modified-site   1276
FT                   /note= "As given in the specification this amino acid can
FT                   also be Leu"
FT   Modified-site   1328
FT                   /note= "As given in the specification this amino acid can
FT                   also be Val"
FT   Modified-site   1452
FT                   /note= "As given in the specification this amino acid can
FT                   also be Tyr"
FT   Modified-site   1472
FT                   /note= "As given in the specification this amino acid can
FT                   also be Ser"
FT   Domain          1640. .2000
FT                   /label= NS4_domain
FT   Modified-site   1877
FT                   /note= "As given in the specification this amino acid can
FT                   also be Gly"
FT   Modified-site   1948
FT                   /note= "As given in the specification this amino acid can
FT                   also be His"
FT   Modified-site   1949
FT                   /note= "As given in the specification this amino acid can
FT                   also be Cys"
FT   Domain          2000. .3011
FT                   /label= NS5_domain
FT   Modified-site   2021
FT                   /note= "As given in the specification this amino acid can
FT                   also be Val"
FT   Modified-site   2348
FT                   /note= "As given in the specification this amino acid can
FT                   also be Ser"
FT   Modified-site   2385
FT                   /note= "As given in the specification this amino acid can
FT                   also be Phe"
FT   Modified-site   2386
FT                   /note= "As given in the specification this amino acid can
FT                   also be Ala"
FT   Modified-site   2502
FT                   /note= "As given in the specification this amino acid can
FT                   also be Phe"
FT   Modified-site   2690
FT                   /note= "As given in the specification this amino acid can
FT                   also be Gly"
FT   Modified-site   2921
FT                   /note= "As given in the specification this amino acid can
FT                   also be Gly"
FT   Modified-site   2996
FT                   /note= "As given in the specification this amino acid can
FT                   also be Pro"
XX
PN   US5712087-A.
XX
PD   27-JAN-1998.
XX
PF   12-MAY-1995;   95US-00440519.
XX
PR   04-APR-1990;   90US-00504352.
PR   07-JUL-1992;   92US-00910760.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kuo G,  Houghton M,  Choo Q;
XX
DR   WPI; 1998-119973/11.
DR   N-PSDB; AAV09989.
XX
PT   Immunoassays for hepatitis C virus antibodies - using combinations of
PT   antigenic fragments of HCV polyprotein.
XX
PS   Disclosure; Fig 1; 59pp; English.
XX
CC   This sequence represents the hepatitis C virus (HCV) polyprotein which is
```

```
CC   used in the construction of novel combinations of HCV antigens that have
CC   a broader range of immunological activity than any single HCV antigen. An
CC   example of such an antigen given in this specification comprises a first
CC   antigen containing at least 8 amino acids of the C domain of the HCV
CC   polyprotein and a second antigen comprising at least 8 amino acids of the
CC   NS3 domain, the NS4 domain, the S domain or the NS5 domain of the HCV
CC   polyprotein in the form of a fusion protein, a physical mixture or bound
CC   to a solid matrix. Note: The features given in the specification as
CC   represented in the feature table of AAW40038 differ from the positions
CC   indicated in Figure 1
XX
SQ   Sequence 3011 AA;

  Query Match             99.5%;  Score 3602;  DB 2;  Length 3011;
  Best Local Similarity   99.6%;  Pred. No. 4.3e-304;
  Matches  683;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 1085

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 1145

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 1205

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 1265

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 1325

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 1385

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
              |||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db       1386 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYT 1445

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 1505

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 1565

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 1625

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 1685

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
              |||||||||||||||||||||||||
Db       1686 VVLSGKPAIIPDREVLYREFDEMEEC 1711


RESULT 14
AAE22049
ID   AAE22049 standard; protein; 3011 AA.
XX
AC   AAE22049;
XX
DT   16-JUL-2002  (first entry)
XX
DE   Hepatitis C virus (HCV) polyprotein.
XX
KW   Hepatitis C virus; HCV; antigen; C domain; polyprotein; NS3 domain;
KW   NS4 domain; S domain; NS5 domain.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .122
FT                   /label= C_domain
FT   Domain          199. .328
FT                   /label= S_domain
FT   Region          1192. .1931
FT                   /note= "c200 polypeptide"
FT   Region          1192. .1457
FT                   /note= "NS3 domain antigen"
FT   Region          1569. .1931
FT                   /note= "NS4 antigen"
FT   Region          2054. .2464
FT                   /note= "NS5 antigen"
XX
PN   US6312889-B1.
XX
PD   06-NOV-2001.
XX
PF   12-MAY-1995;   95US-00440549.
XX
PR   04-APR-1990;   90US-00504352.
PR   07-JUL-1992;   92US-00910760.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Houghton M,  Choo Q,  Kuo G;
XX
DR   WPI; 2002-040268/05.
DR   N-PSDB; AAD35043.
XX
PT   Combination of hepatitis C viral (HCV) antigens, useful in improved
PT   immunoassay for detecting HCV antibodies.
XX
PS   Example 1; Col 45-60; 58pp; English.
XX
CC   The invention relates to combination of hepatitis C viral (HCV) antigens
CC   that have a broader range of immunological reactivity than any single HCV
CC   antigen. The combinations consist of an antigen from the C domain of the
CC   HCV polyprotein, and at least one additional HCV antigen from either the
CC   NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
CC   the form of fusion protein, a simple physical mixture, or the individual
CC   antigens commonly bound to a solid matrix. The combinations of antigens
CC   provides broad range immunoassays for anti-HCV antibodies. The invention
CC   therefore provides a method for detecting antibodies to HCV in a mammal
CC   suspected of containing such antibodies. The present sequence is HCV
CC   polyprotein. Note: This sequence SEQ.ID.NO:10 is stated to be similar to
CC   the sequence shown in Fig 1 (AAE22052) of the specification. However
CC   these sequences differ
XX
SQ   Sequence 3011 AA;

  Query Match             99.5%;  Score 3602;  DB 5;  Length 3011;
  Best Local Similarity   99.6%;  Pred. No. 4.3e-304;
  Matches  683;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 1085

Qy         61 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 GTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 1145

Qy        121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1146 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 1205

Qy        181 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 RSPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFG 1265

Qy        241 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 AYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATS 1325

Qy        301 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 ILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVI 1385

Qy        361 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYT 420
              ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db       1386 KGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVATDALMTGYT 1445

Qy        421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 GDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPG 1505

Qy        481 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 ERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGL 1565

Qy        541 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 THIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLL 1625

Qy        601 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 YRLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGR 1685

Qy        661 VVLSGKPAIIPDREVLYREFDEMEEC 686
              |||||||||||||||||||||||||
Db       1686 VVLSGKPAIIPDREVLYREFDEMEEC 1711
```

RESULT 15
AAE18688
ID AAE18688 standard; protein; 728 AA.
XX
AC AAE18688;
XX
DT 17-MAY-2002 (first entry)
XX
DE NS3/4a mutant conformational antigen.
XX
KW Hepatitis C virus; NS3/4a antigen; HCV infection; mutant; mutein.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 182
FT /note= "Wild type Ser is substituted with Ala"
XX
PN WO200196875-A2.
XX
PD 20-DEC-2001.
XX
PF 14-JUN-2001; 2001WO-US019369.
XX
PR 15-JUN-2000; 2000US-0212082P.
PR 02-APR-2001; 2001US-0280811P.
PR 02-APR-2001; 2001US-0280867P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Chien DY, Arcangel P, Tandeske L, George-Nasciemento C, Coit D;
PI Medina-Selby A;
XX
DR WPI; 2002-179522/23.
XX
PT Immunoassay solid support useful for detecting hepatitis C virus
PT infection in a biological sample, comprises at least one of HCV anti-core
PT antibody and HCV NS3/4a epitope, bound to the support.
XX
PS Disclosure; Fig 3; 87pp; English.
XX
CC The present invention relates to hepatitis C virus (HCV) core antigen and
CC NS (nonstructural) 3/4a antibody combination assay that can detect both
CC HCV antigens and antibodies present in a sample using a single solid
CC matrix as well as immunoassay solid supports for use in the assay. The
CC solid support is useful for detecting HCV infection in a biological
CC sample. The present sequence is NS3/4a mutant conformational antigen.
CC This sequence is used in the exemplification of the invention
XX
SQ Sequence 728 AA;

Query Match 99.5%; Score 3600; DB 5; Length 728;
Best Local Similarity 99.6%; Pred. No. 7.9e-305;
Matches 682; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         44 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 103

Qy         62 TRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG 121
              |||||||||||||||||||||||||||||| ||:||||||||||||||||||||||||||
Db        104 TRTIASPKGPVIQMYTNVDQDLVGWPASQGTRSLTPCTCGSSDLYLVTRHADVIPVRRRG 163

Qy        122 DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMR 181
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        164 DSRGSLLSPRPISYLKGSAGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMR 223

Qy        182 SPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGA 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 SPVFTDNSSPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGA 283

Qy        242 YMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 YMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSI 343

Qy        302 LGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 LGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEVIK 403

Qy        362 GGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYTG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 GGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPPIGDVVVVATDALMTGYTG 463

Qy        422 DFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPGE 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 DFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQDAVSRTQRRGRTGRGKPGIYRFVAPGE 523

Qy        482 RPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGLT 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 RPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTGLT 583

Qy        542 HIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLY 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 HIDAHFLSQTKQSGENLPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLY 643

Qy        602 RLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRV 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 RLGAVQNEITLTHPVTKYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRV 703

Qy        662 VLSGKPAIIPDREVLYREFDEMEEC 686
              |||||||||||||||||||||||||
Db        704 VLSGKPAIIPDREVLYREFDEMEEC 728
```

Search completed: June 21, 2004, 10:30:03
Job time : 44.9396 secs